Schaum

Probabilidad y estadística

Cuarta edición

897 problemas totalmente resueltos

20 videos en línea de problemas resueltos

UTILÍCELO EN LAS SIGUIENTES ASIGNATURAS

INTRODUCCIÓN A LA PROBABILIDAD Y LA ESTADÍSTICA

PROBABILIDAD • ESTADÍSTICA EN LOS NEGOCIOS

ESTADÍSTICA BÁSICA • PRINCIPIOS DE ESTADÍSTICA

Murray R. Spiegel • John Schiller • R. Alu Srinivasan

Mc Graw Hill Education

Probabilidad y estadística

Los siguientes profesores participaron en la revisión de la presente obra y aportaron ejercicios, que fueron anexados al final de cada capítulo en el apartado “Ejercicios aportados”. Agradecemos esta valiosa colaboración por parte de dichos revisores.

Domingo Alarcón Ortiz
Universidad Nacional Autónoma de México

Alfonso Reyes Martínez
Centro de Bachillerato Tecnológico Industrial y de Servicios # 89

Andrés Basilio Ramírez y Villa
Universidad Nacional Autónoma de México

Antonino Pérez Hernández
Centro de Investigación en Materiales Avanzados, S. C.

Daniel Sepúlveda Jiménez
Universidad Autónoma de Chapingo

Efraín García Venegas
Instituto Tecnológico de Celaya

Samuel Rosalio Cuevas
Universidad de Guadalajara

Eva Valdez Alemán
Instituto Politécnico Nacional

Fermín Acosta Magallanes
Instituto Politécnico Nacional

Germán Caudana Camacho
Universidad de Sonora, Campus Hermosillo

Gustavo Ernesto Guridi Ramos
Universidad Veracruzana, Campus Veracruz

Ignacio Fonseca Chon
Universidad de Sonora

Irene Patricia Valdez y Alfaro
Universidad Nacional Autónoma de México

Israel Portillo Arroyo
Instituto Tecnológico de Parral, Chihuahua

Jorge Velázquez Centeno
Universidad Iberoamericana, Campus León

José del Carmen Jiménez Hernández
Universidad Tecnológica de la Mixteca

Ismael Lara García
Universidad Veracruzana, Campus Veracruz

Luz María Trinidad Pérez Alvarado
Instituto Tecnológico de León

María del Rosario Lara Delfín
Instituto Tecnológico de Toluca

Marcelo Romero Huertas
Universidad Autónoma del Estado de México

Miguel Casas Ramírez
Instituto Tecnológico del Valle del Yaqui

Miguel de Nazareth Pineda Becerril
Universidad Nacional Autónoma de México, Campus Cuautitlán

Rafael Gustavo Alfaro Pérez
Instituto Tecnológico de Matamoros

Rocío Cerecero López
Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Cuernavaca

Virgilio Requena Guerrero
Universidad Autónoma de Tamaulipas

Rosa María Rodríguez González
Universidad Iberoamericana, Campus León

Ramón Sebastián Salat Figols
Instituto Politécnico Nacional

Rubén Vinagre Reyes
Universidad Juárez Autónoma de Tabasco

PROBABILIDAD Y ESTADÍSTICA

Cuarta edición

MURRAY R. SPIEGEL
Ex profesor y coordinador de Matemáticas
Rensselaer Polytechnic Institute
Hartford Graduate Center

John J. Schiller
Profesor asociado de Matemáticas
Temple University

R. Alu Srinivasan
Profesor de Matemáticas
Temple University

Revisión técnica:

Alejandra Vargas Espinoza de los Monteros
Universidad Nacional Autónoma de México

MÉXICO • BOGOTÁ • BUENOS AIRES • CARACAS • GUATEMALA • MADRID • NUEVA YORK
SAN JUAN • SANTIAGO • SAO PAULO • AUCKLAND • LONDRES • MILÁN • MONTREAL
NUEVA DELHI • SAN FRANCISCO • SAO PAULO • SINGAPUR • ST. LOUIS • SIDNEY • TORONTO

Director general México: Miguel Ángel Toledo Castellanos
Editor sponsor: Jesús Mares Chacón
Coordinadora editorial: Marcela I. Rocha Martínez
Editor de desarrollo: Edmundo Carlos Zúñiga Gutiérrez
Supervisor de producción: Zeferino García García
Traductor: Gabriel Nagore Cázares

PROBABILIDAD Y ESTADÍSTICA
Cuarta edición

Prohibida la reproducción total o parcial de esta obra,
por cualquier medio, sin la autorización escrita del editor.

DERECHOS RESERVADOS © 2015 respecto a la prımera edición en español por
McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V.
Prolongación Paseo de la Reforma 1015, Torre A,
Piso 17, Colonia Desarrollo Santa Fe,
Delegación Álvaro Obregón,
C.P. 01376, México, D.F.
Miembro de la Cámara Nacional de la Industria Editorial Mexicana, Reg. Núm. 736

ISBN: 978-1-4562-4002-8

Obra original: *Probabilidad y estadística* © 2013,
respecto a la cuarta edición en español, por McGraw-Hill Interamericana Editores, S. A. de C. V.

ISBN edición original: 978-607-15-1188-1
ISBN (edición anterior): 978-607-15-0270-4

Traducido de la cuarta edición de: Schaum's Outline of *Probability and Statistics,* de Murray R. Spiegel, John J. Schiller y R. Alu Srinivasan. Copyright © 2013, by The McGraw-Hill/Companies, Inc. All rights reserved.
978-0-07-179557-9

ANR 01/14

1234567890 2356789014

PREFACIO

El importante y fascinante tema de la probabilidad comenzó en el siglo XVII, con los trabajos de matemáticos como Fermat y Pascal, para responder a preguntas referentes a juegos de azar. No fue sino hasta el siglo XX que se desarrolló una teoría matemática rigurosa basada en axiomas, definiciones y teoremas. Con el tiempo, la teoría de la probabilidad encontró muchas aplicaciones, no sólo en la ingeniería, la ciencia y las matemáticas, sino también en campos que van desde la ciencia actuarial, la agricultura y los negocios hasta áreas como la medicina y la psicología. En muchos casos las mismas aplicaciones contribuyeron al ulterior desarrollo de esta teoría.

La estadística, que se originó mucho antes que la probabilidad, se ocupaba, de manera principal, de la recolección, organización y presentación de datos en tablas y gráficas. Con el advenimiento de la probabilidad se comprobó que la estadística podía ser utilizada para obtener conclusiones válidas y tomar decisiones razonables con base en el análisis de datos, por ejemplo en la teoría del muestreo y en la predicción o pronóstico.

El propósito del presente libro es presentar una introducción moderna a la probabilidad y a la estadística usando los conocimientos del cálculo. La obra se divide en dos partes: la primera se ocupa de la probabilidad (y puede utilizarse como introducción al tema), mientras que la segunda aborda la estadística.

Desde la concepción de esta obra, se pensó en que sirviera como libro de texto en un curso formal de probabilidad y estadística o como suplemento de los textos estándares actuales. También será de valor considerable como libro de referencia para investigadores o para aquellos autodidactas interesados en este campo. El libro puede utilizarse para un curso de un año, o, mediante una cuidadosa elección de los temas, para uno de un semestre.

En esta cuarta edición se han incluido problemas nuevos, propuestos por profesores de América Latina. Estos usuarios también participaron como revisores técnicos de la presente edición en español. Sus nombres figuran en la portadilla de la obra y sus aportaciones se encuentran al final de cada capítulo, en el apartado "Problemas aportados". Les estamos enormemente agradecidos por su participación.

Desde la edición anterior, además de las mejoras en todo el libro, se agregó un capítulo sobre estadística no paramétrica para ampliar la aplicabilidad del texto sin elevar su nivel. Este tema continúa en esta edición, en la que se ha reformado el libro y se ha agregado un capítulo sobre métodos bayesianos. En los últimos años, el paradigma bayesiano ha ganado renombre y un efecto crecientes en áreas tales como la economía, las ciencias ambientales, la medicina y las finanzas. Como el análisis estadístico bayesiano requiere de muchos cálculos, los avances de la informática le han permitido ganar mayor aceptación. Por lo tanto, es imprescindible una introducción a los principios básicos del análisis de datos bayesiano, lo cual además es coherente con el propósito principal del profesor Murray R. Spiegel cuando escribió la obra original: presentar una introducción moderna a la probabilidad y a la estadística usando los conocimientos del cálculo.

La primera edición de *Probabilidad y estadística* de la serie Schaum, de Murray R. Spiegel, apareció en 1975, y desde entonces se ha reimpreso en 21 ocasiones. Su primo cercano, *Estadística,* del mismo autor y también de la serie Schaum, fue descrito por Gian-Carlo Rota en su libro *Indiscrete Thoughts* como la más clara introducción a la estadística, por lo que se emprendió esta revisión con profundo respeto y cierta precaución. Nuestro principio guía era realizar cambios sólo donde se hiciera necesario para poner esta obra a tono con los temas más importantes que incluyen los libros contemporáneos. El amplio estudio de los conjuntos, material introductorio estándar en los libros de las décadas de 1960 y 1970, se redujo de manera considerable. La definición de variable aleatoria continua es ahora el estándar; se da más importancia a la función de distribución acumulada en razón de que es un concepto fundamental de mayor importancia que la densidad de probabilidad. También se asigna más importancia a los valores P de las pruebas de hipótesis, puesto que la tecnología ha hecho posible la fácil determinación de estos valores, que proporcionan una información más específica que sólo determinar si una prueba satisface o no un determinado nivel de significancia especificado de antemano. La tecnología también ha permitido eliminar las tablas logarítmicas. Un capítulo sobre estadística no paramétrica ha sido agregado para ampliar la aplicabilidad de este libro sin elevar su nivel. Algunos conjuntos de problema han sido eliminados, en especial en los casos en que se trataba de pruebas de teoremas para las que no se daba ninguna pista o ayuda de ninguna clase. Consideramos que, en general, se ha preservado el propósito principal de la primera edición: presentar una introducción moderna a la probabilidad y a la estadística usando los antecedentes del cálculo, así como las características que hicieron que esa primera edición tuviera un éxito tan grande. Esperamos que esta edición pueda servir a una gama aún más amplia de estudiantes.

AGRADECIMIENTO

Agradecemos al albacea literario del desaparecido Sir Ronald A. Fisher, F.R.S., al Dr. Frank Yates, F.R.S. y al Grupo Longman Ltd., Londres, por permitirme utilizar la tabla III de su libro *Statistical Tables for Biological, Agricultural and Medical Research* (6a. edición, 1974). También deseo aprovechar esta oportunidad para agradecer a David Beckwith por su excepcional corrección del texto, y a Nicola Monti por sus excelentes ilustraciones.

J. SCHILLER
R. A. SRINIVASAN

CONTENIDO

Parte I PROBABILIDAD 1

CAPÍTULO 1 Probabilidad básica 3

Experimentos aleatorios 3
Espacios muestrales 3
Eventos 4
Concepto de probabilidad 5
Axiomas de la probabilidad 5
Algunos teoremas importantes acerca de la probabilidad 5
Asignación de probabilidades 6
Probabilidad condicional 7
Teoremas de probabilidad condicional 7
Eventos independientes 7
Teorema o regla de Bayes 8
Análisis combinatorio 8
Principio fundamental de conteo: diagramas de árbol 8
Permutaciones 9
Combinaciones 9
Coeficiente binomial 10
Aproximación de Stirling para $n!$ 10

CAPÍTULO 2 Variables aleatorias y distribuciones de probabilidad 34

Variables aleatorias 34
Distribuciones de probabilidad discretas 34
Funciones de distribución de variables aleatorias 35
Funciones de distribución de variables aleatorias discretas 35
Variables aleatorias continuas 36
Interpretaciones gráficas 38
Distribuciones conjuntas 39
Variables aleatorias independientes. 41
Cambio de variables 41
Distribuciones de probabilidad de funciones de variables aleatorias 42
Convoluciones 43
Distribuciones condicionales 43
Aplicaciones a la probabilidad geométrica 44

CAPÍTULO 3 Esperanza matemática 75

Definición de esperanza matemática 75
Funciones de variables aleatorias 76
Algunos teoremas sobre la esperanza 76
La varianza y la desviación estándar 77
Algunos teoremas sobre la varianza 78
Variables aleatorias estandarizadas 78
Momentos 78
Funciones generadoras de momentos 79
Algunos teoremas sobre funciones generadoras de momentos 79
Funciones características 80
Varianza de distribuciones conjuntas. Covarianza 81
Coeficiente de correlación 82
Esperanza, varianza y momentos condicionales 82
Desigualdad de Chebyshev 83
Ley de los grandes números 83
Otras medidas de tendencia central 83
Percentiles 84
Otras medidas de dispersión 84
Sesgo y curtosis 84

CAPÍTULO 4 Distribuciones especiales de probabilidad 108

La distribución binomial 108
Algunas propiedades de la distribución binomial 108
Ley de los grandes números para ensayos de Bernoulli 109
Distribución normal 109
Algunas propiedades de la distribución normal 110
Relación entre las distribuciones binomial y normal 111
Distribución de Poisson 111
Algunas propiedades de la distribución de Poisson 111
Relación entre las distribuciones binomial y de Poisson 111
Relación entre las distribuciones de Poisson y la normal 112
Teorema del límite central 112
Distribución multinomial 112
Distribución hipergeométrica 112
Distribución uniforme 113
Distribución de Cauchy 114
Distribución gamma 114
Distribución beta 114
Distribución ji cuadrada 115
Distribución *t* de Student 115
Distribución *F* 116
Relación entre las distribuciones ji cuadrada, *t* y *F* 117
Distribución normal bivariada 117
Diversas distribuciones 117

Parte II **ESTADÍSTICA** 151

CAPÍTULO 5 Teoría del muestreo 153

Población y muestra. Inferencia estadística 153
Muestreo con y sin reemplazo 153
Muestras aleatorias. Números aleatorios. 154
Parámetros poblacionales 154
Estadísticos muestrales 154
Distribuciones muestrales 155
Media muestral 155
Distribución muestral de medias 155
Distribución muestral de proporciones 156
Distribución muestral de diferencias y sumas. 157
Varianza muestral 157
Distribución muestral de las varianzas 158
Caso en el que no se conoce la varianza poblacional 159
Distribución muestral de razones de varianzas 159
Otros estadísticos 159
Distribuciones de frecuencias 160
Distribución de frecuencias relativas 161
Cálculo de la media, varianza y momentos para datos agrupados 161

CAPÍTULO 6 Teoría de la estimación 195

Estimadores insesgados y eficientes 195
Estimaciones puntuales y por intervalos: confiabilidad 195
Estimaciones del intervalo de confianza de parámetros poblacionales. 195
Intervalos de confianza para medias 196
Intervalo de confianza para proporciones 197
Intervalos de confianza para diferencias y sumas 197
Intervalos de confianza para la varianza de una distribución normal 197
Intervalos de confianza para razones, o cocientes, de varianza. 198
Estimaciones de máxima verosimilitud 198

CAPÍTULO 7 Pruebas de hipótesis y significancia 213

Decisiones estadísticas 213
Hipótesis estadísticas. Hipótesis nulas 213
Pruebas de hipótesis y significancia 213
Errores del tipo I y del tipo II 213
Nivel de significancia 214
Pruebas en las que interviene la distribución normal 214
Pruebas de una cola y de dos colas 215
Valor *P* 215
Pruebas especiales de significancia en el caso de muestras grandes. 216
Pruebas especiales de significancia de muestras pequeñas 217
Relación entre la teoría de la estimación y la prueba de hipótesis 219

Curvas características de operación. Potencia de una prueba 219
Gráficas de control de calidad 219
Ajuste de distribuciones teóricas a distribuciones de frecuencia muestrales 219
La prueba ji cuadrada para la bondad de ajuste 220
Tablas de contingencia 221
Corrección de Yates para la continuidad 221
Coeficiente de contingencia 222

CAPÍTULO 8 Ajuste de curvas, regresión y correlación 265

Ajuste de curvas 265
Regresión. 265
Método de mínimos cuadrados 266
Recta de mínimos cuadrados. 267
Recta de mínimos cuadrados en términos de varianzas y covarianza muestrales. . . . 268
Parábola de mínimos cuadrados 268
Regresión múltiple 269
Error estándar de estimaciones 269
Coeficiente de correlación lineal. 270
Coeficiente de correlación generalizada 271
Correlación de rangos 271
Interpretación probabilística de la regresión. 272
Interpretación probabilística de correlación 273
Teoría muestral de la regresión 273
Teoría muestral de correlación 274
Correlación y dependencia 274

CAPÍTULO 9 Análisis de varianza. 314

El propósito del análisis de varianza. 314
Clasificación unidireccional o experimentos de un factor 314
Variación total. Variación dentro de los tratamientos. Variación entre tratamientos 315
Métodos cortos para obtener variaciones 315
Modelo matemático lineal para análisis de varianza. 315
Valores esperados de las variaciones. 316
Distribuciones de las variaciones 317
La prueba F para la hipótesis nula de medias iguales. 317
Análisis de tablas de varianza 317
Modificaciones para números desiguales de observaciones 318
Clasificación bidireccional o experimentos de dos factores 318
Notación para experimentos de dos factores. 318
Variaciones para experimentos de dos factores. 319
Análisis de la varianza para experimentos de dos factores 320
Experimentos de dos factores con replicación 321
Diseño experimental 323

CAPÍTULO 10 Pruebas no paramétricas 348

Introducción 348
Prueba de los signos 348
Prueba *U* de Mann-Whitney 349
Prueba *H* de Kruskal-Wallis 350
Prueba *H* corregida para empates 350
Prueba de corridas (rachas) de aleatoriedad 350
Aplicaciones adicionales para la prueba de corridas. 351
Correlación de rangos de Spearman 352

CAPÍTULO 11 Métodos bayesianos 372

Probabilidad subjetiva 372
Distribuciones a priori y a posteriori. 372
Muestreo a partir de una población binomial 375
Muestreo a partir de una población de Poisson. 376
Muestreo a partir de una población normal con varianza conocida 377
Distribuciones a priori impropias 378
Distribuciones a priori conjugadas 379
Estimación puntual bayesiana. 380
Estimación del intervalo bayesiano. 382
Pruebas de hipótesis bayesianas 383
Factores de Bayes 384
Distribuciones predictivas bayesianas. 386

APÉNDICE A Temas matemáticos 412

APÉNDICE B Ordenadas *y* de la curva normal estándar en *z* 415

APÉNDICE C Áreas bajo la curva normal estándar de 0 a *z* 416

APÉNDICE D Valores del percentil t_p para la distribución *t* de Student con *v* grados de libertad 417

APÉNDICE E Valores del percentil χ^2_p para la distribución ji cuadrada con *v* grados de libertad 418

APÉNDICE F Valores del percentil 95 (niveles de 0.05), $F_{0.95}$ para la distribución *F* 419

APÉNDICE G Valores de $e^{-\lambda}$ 421

APÉNDICE H Números aleatorios 421

ÍNDICE ANALÍTICO 422

ÍNDICE DE PROBLEMAS RESUELTOS 426

Parte I

CAPÍTULO 1

Probabilidad básica

EXPERIMENTOS ALEATORIOS

Para todos es conocida la importancia de los experimentos en la ciencia y en la ingeniería. La experimentación es útil porque permite suponer que si se realiza un determinado experimento bajo condiciones esencialmente idénticas, se llegará a resultados básicamente iguales. En estas circunstancias, puede controlarse el valor de las variables que afectan el resultado de un experimento.

Sin embargo, en algunos experimentos no es posible conocer o controlar el valor de ciertas variables, de modo que los resultados variarán de una a otra realización del experimento, aun cuando la mayor parte de las condiciones sean iguales. Estos experimentos se describen como *aleatorios*. Los siguientes son algunos ejemplos.

EJEMPLO 1.1 Si se lanza una moneda, el resultado del experimento será que caiga "cruz", simbolizada por T (o bien 0), o que tal vez caiga "cara", simbolizada por H (o bien 1), es decir, por uno de los elementos del conjunto $\{H, T\}$ (o bien del conjunto $\{0, 1\}$).

EJEMPLO 1.2 Si se lanza un dado, el resultado del experimento es que caiga uno de los números del conjunto: {l, 2, 3, 4, 5, 6}.

EJEMPLO 1.3 Si se lanza una moneda dos veces, habrá cuatro resultados posibles, que se indican por $\{HH, HT, TH, TT\}$, es decir, dos caras, una cara en el primer lanzamiento y una cruz en el segundo, etcétera.

EJEMPLO 1.4 Si se fabrican pernos con una máquina, como resultado de este experimento se tendrá alguno defectuoso. Así, cuando se hace un perno, éste será un miembro del conjunto {defectuoso, no defectuoso}.

EJEMPLO 1.5 Si un experimento consiste en medir el "tiempo de vida" de los focos de luz eléctrica producidos por una empresa, entonces el resultado del experimento será un tiempo t en horas, que se encuentra en algún intervalo —por ejemplo, $0 \leq t \leq 4\,000$—, donde se supone que ningún foco dura más de 4 000 horas.

ESPACIOS MUESTRALES

A un conjunto S que consta de todos los resultados posibles de un experimento aleatorio se le llama *espacio muestral*, y a cada resultado se le llama *punto muestral*. A menudo habrá más de un espacio muestral que describa los resultados de un experimento, pero, en general, solamente habrá uno que proporcione la mayor información.

EJEMPLO 1.6 Si se lanza un dado, un espacio muestral, o conjunto de todos los resultados posibles, será {1, 2, 3, 4, 5, 6}, mientras que otro será {par, impar}. Sin embargo, es claro que el último no será adecuado para determinar, por ejemplo, si un resultado es divisible entre 3.

A menudo es útil representar gráficamente un espacio muestral. En tales casos, siempre que sea posible, es deseable utilizar números en lugar de letras.

EJEMPLO 1.7 Si se lanza una moneda dos veces y se utiliza 0 para representar cruz y 1 para representar cara, el espacio muestral (vea el ejemplo 1.3) puede representarse por puntos como en la figura 1-1, donde, por ejemplo, (0, 1) representa cruz en el primer lanzamiento y cara en el segundo lanzamiento, es decir, TH.

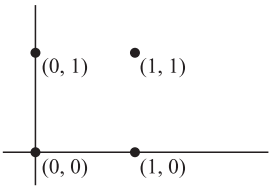


Figura 1-1

Si un espacio muestral tiene un número finito de puntos, como en el ejemplo 1.7, se llama *espacio muestral finito*. Si tiene tantos puntos como números naturales 1, 2, 3,…, se llama *espacio muestral contable infinito*. Si tiene tantos puntos como hay en un cierto intervalo del eje x, como en $0 \leq x \leq 1$, se llama *espacio muestral incontable infinito*. A un espacio muestral que es contable finito o infinito, a menudo se le llama *espacio muestral discreto*, mientras que a uno que es incontable infinito se le denomina *espacio muestral no discreto* o *continuo.*

EVENTOS

Un evento, o suceso, es un subconjunto A del espacio muestral S, es decir, también es un conjunto de resultados posibles. Si el resultado de un experimento es un elemento de A, se dice que *ha ocurrido* el evento A. Un evento que consta de un solo punto de S suele llamarse *evento simple* o *elemental*.

EJEMPLO 1.8 Si se lanza una moneda dos veces, el evento en el que sólo cae una cara es el subconjunto del espacio muestral que consta de los puntos (0, 1) y (1, 0), como se ilustra en la figura 1-2.

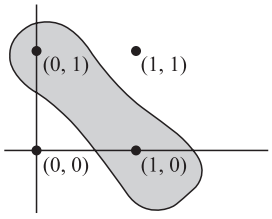


Figura 1-2

Como eventos particulares se tiene al mismo S, que es el *evento seguro* o *cierto*, puesto que debe ocurrir algún elemento de S, y el conjunto vacío $\varnothing$, que se llama *evento imposible* porque no puede ocurrir un elemento del $\varnothing$.

Empleando las operaciones de conjuntos con los eventos de S, pueden obtenerse otros eventos de S. Por ejemplo, si A y B son eventos, entonces

1. $A \cup B$ es el evento "A o B o ambos". $A \cup B$ se llama la *unión* de A y B.
2. $A \cap B$ es el evento "A y B". $A \cap B$ se denomina la *intersección* de A y B.
3. A' es el evento "no A". A' se llama el *complemento* de A.
4. $A - B = A \cap B'$ es el evento "A pero no B". En particular, $A' = S - A$.

Si los conjuntos que corresponden a los eventos A y B son ajenos, o disjuntos, es decir, $A \cap B = \varnothing$, se suele decir que los eventos son *mutuamente excluyentes*. Esto significa que no pueden ocurrir simultáneamente los dos. Se dice que una colección de los eventos $A_1, A_2, \ldots, A_n$ es mutuamente excluyente si cada par de eventos de la colección es mutuamente excluyente.

EJEMPLO 1.9 Volviendo al experimento de lanzar una moneda dos veces, sea A el evento "que caiga por lo menos una cara" y B el evento "el segundo lanzamiento que caiga cruz". Entonces, $A = \{HT, TH, HH\}$, $B = \{HT, TT\}$ y, por tanto, se tiene:

$$A \cup B = \{HT, TH, HH, TT\} = S \qquad A \cap B = \{HT\}$$
$$A' = \{TT\} \qquad A - B = \{TH, HH\}$$

CONCEPTO DE PROBABILIDAD

En cualquier experimento aleatorio existe siempre incertidumbre respecto a si un determinado acontecimiento ocurrirá o no. Como una medida de la *posibilidad*, o *probabilidad*, con la cual puede esperarse que ocurra el evento, suele emplearse un número entre 0 y 1. Si se está seguro de que ocurra un determinado evento, se dice que su probabilidad es de 100%, o bien 1, pero si se está seguro de que no ocurra, se dice que su probabilidad es cero. Si, por ejemplo, la probabilidad es $\frac{1}{4}$, se dice que hay una posibilidad de 25% de que ocurra y una posibilidad de 75% de que no ocurrirá. De manera equivalente, puede decirse que las *posibilidades* en contra de su ocurrencia son de 75% contra 25%, o bien 3 a 1.

Hay dos enfoques importantes por medio de los cuales puede estimarse la probabilidad de un evento.

1. **ENFOQUE CLÁSICO.** Si un evento puede ocurrir de h diferentes maneras de un total de n maneras posibles, igualmente probables, entonces la probabilidad de ese evento es h/n.

EJEMPLO 1.10 Suponga que se desea saber cuál es la probabilidad de que en un solo lanzamiento de una moneda se obtenga una cara. Puesto que una moneda puede caer de dos maneras igualmente probables, cara o cruz (se supone que la moneda no rueda y se pierde, o queda parada sobre su borde), y que de estas dos posibilidades sólo se puede obtener una cara, se concluirá que la probabilidad estimada es $\frac{1}{2}$. Para llegar a esto, se supone que la moneda es *legal*, es decir, que no está *cargada*.

2. **ENFOQUE DE FRECUENCIA.** Si después de repetir n veces un experimento, donde n es muy grande, se observa que un evento ocurre en h de estas veces, entonces la probabilidad del evento es h/n. A esto se le llama *probabilidad empírica* del evento.

EJEMPLO 1.11 Si se lanza una moneda 1 000 veces y se encuentra que en 532 de estos lanzamientos cae cara, se estima que la probabilidad de obtener una cara es $532/1\,000 = 0.532$.

Tanto el enfoque clásico como el de frecuencia tienen inconvenientes serios, el primero porque la expresión "igualmente probables" es vaga y el segundo porque la "cantidad grande" es vago. Debido a estas dificultades, los matemáticos han llegado a un *enfoque axiomático* para la probabilidad.

AXIOMAS DE LA PROBABILIDAD

Suponga que se tiene un espacio muestral S. Si S es discreto, todos los subconjuntos corresponden a eventos y recíprocamente, pero si S no es discreto, sólo los subconjuntos especiales (llamados medibles) corresponden a eventos. A cada evento A de una clase C de eventos se le asocia un número real $P(A)$. Entonces P se llama *función de probabilidad*, y $P(A)$ es la *probabilidad* del evento A, si se satisfacen los axiomas siguientes:

Axioma 1 Para cada evento A de la clase C,

$$P(A) \geq 0 \tag{1}$$

Axioma 2 Para el evento cierto o seguro S de la clase C,

$$P(S) = 1 \tag{2}$$

Axioma 3 Para cualquier número de eventos mutuamente excluyentes $A_1, A_2, \ldots$, de la clase C,

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots \tag{3}$$

En particular, dados dos eventos mutuamente excluyentes A_1, A_2,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \tag{4}$$

ALGUNOS TEOREMAS IMPORTANTES ACERCA DE LA PROBABILIDAD

De acuerdo con los axiomas anteriores, pueden demostrarse varios teoremas acerca de la probabilidad que son importantes en el trabajo subsiguiente.

Teorema 1-1 Si $A_1 \cup A_2$, entonces $P(A_1) \leq P(A_2)$ y $P(A_2 - A_1) = P(A_2) - P(A_1)$.

Teorema 1-2 Para todo evento A,

$$0 \leq P(A) \leq 1, \tag{5}$$

es decir, una probabilidad está entre 0 y 1.

Teorema 1-3 $$P(\varnothing) = 0 \tag{6}$$

es decir, el evento imposible tiene probabilidad cero.

Teorema 1-4 Si A' es el complemento de A, entonces

$$P(A') = 1 - P(A) \tag{7}$$

Teorema 1-5 Si $A = A_1 \cup A_2 \cup \cdots \cup A_n$ donde $A_1, A_2, \ldots A_n$ son eventos mutuamente excluyentes, entonces

$$P(A) = P(A_1) + P(A_2) + \cdots + P(A_n) \tag{8}$$

En particular, si $A = S$, el espacio muestral, entonces

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1 \tag{9}$$

Teorema 1-6 Si A y B son dos eventos, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{10}$$

De manera más general, si A_1, A_2, A_3 son tres eventos cualesquiera, entonces

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = {} & P(A_1) + P(A_2) + P(A_3) \\ & - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_3 \cap A_1) \\ & + P(A_1 \cap A_2 \cap A_3) \end{aligned} \tag{11}$$

Este teorema también puede generalizarse a n eventos.

Teorema 1-7 Para todo par de eventos A y B,

$$P(A) = P(A \cap B) + P(A \cap B') \tag{12}$$

Teorema 1-8 Si un evento A debe ser el resultado de la ocurrencia de uno de los eventos mutuamente excluyente $A_1, A_2, \ldots, A_n$, entonces

$$P(A) = P(A \cap A_1) + P(A \cap A_2) + \cdots + P(A \cap A_n) \tag{13}$$

ASIGNACIÓN DE PROBABILIDADES

Si un espacio muestral S consta de un número finito de resultados $a_1, a_2, \ldots, a_n$, entonces, de acuerdo con el teorema 1-5,

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1 \tag{14}$$

donde, $A_1, A_2, \ldots, A_n$ son eventos elementales dados por $A_i = \{a_i\}$.

De lo anterior se deduce que, como probabilidades de estos eventos simples, se pueden elegir, de manera arbitraria, números no negativos cualesquiera siempre y cuando se satisfaga (14). En particular, si se suponen *probabilidades iguales* para todos los eventos simples, entonces

$$P(A_k) = \frac{1}{n}, \quad k = 1, 2, \ldots, n \tag{15}$$

y si A es cualquier evento compuesto de h eventos simples, se tiene

$$P(A) = \frac{h}{n} \tag{16}$$

Esto es equivalente al enfoque clásico de probabilidad dado en la página 5. Por supuesto, pueden utilizarse otros procedimientos para asignar probabilidades, como, por ejemplo, el método de la frecuencia de la página 5.

La asignación de probabilidades proporciona un *modelo matemático*, cuyo éxito debe probarse mediante experimentación, de la misma manera que se prueban mediante experimentación las teorías de la física o de otras ciencias.

EJEMPLO 1.12 Se lanza un dado una sola vez. Encuentre la probabilidad de que se obtenga un 2 o bien un 5.

El espacio muestral es $S = \{1, 2, 3, 4, 5, 6\}$. Si a los puntos de la muestra se les asignan probabilidades iguales, es decir, si se supone que el dado no está cargado, entonces

$$P(1) = P(2) = \cdots = P(6) = \frac{1}{6}$$

El evento que se obtenga, un 2 o un 5, se representa como $\{2\} \cup \{5\}$. Por tanto,

$$P(\{2\} \cup \{5\}) = P(2) + P(5) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

PROBABILIDAD CONDICIONAL

Sean A y B dos eventos (figura 1-3) tales que $P(A) > 0$. Denote mediante $P(B|A)$ la probabilidad de B *dado que* A ha ocurrido. Como se sabe que A ha ocurrido, se convierte en el nuevo espacio muestral que sustituye al original S. Esto conduce a la definición

$$P(B|A) \equiv \frac{P(A \cap B)}{P(A)} \tag{17}$$

o bien

$$P(A \cap B) \equiv P(A)\,P(B|A) \tag{18}$$

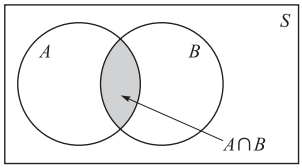


Figura 1-3

Expresado en palabras, (18) indica que la probabilidad de que ocurran tanto A como B es igual a la posibilidad de que ocurra A, por la probabilidad de que ocurra B, dado que ha ocurrido A. A $P(B|A)$ se le llama *probabilidad condicional* de B dado A, es decir, la probabilidad de que ocurra B dado que ha ocurrido A. Es fácil demostrar que la probabilidad condicional satisface los axiomas dados en la página 5.

EJEMPLO 1.13 Encuentre la probabilidad de que en un solo lanzamiento de un dado se obtenga un número menor que 4 si *a*) no se da ninguna otra información y *b*) se sabe que en ese lanzamiento se obtuvo un número impar.

a) Sea B el suceso {menor que 4}. Puesto que B es la unión de los eventos 1, 2 o 3, se tiene, de acuerdo con el teorema 1-5,

$$P(B) = P(1) + P(2) + P(3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

suponiendo probabilidades iguales para los puntos de la muestra.

b) Sea A el evento {número impar}, se ve que $P(A) = \frac{3}{6} = \frac{1}{2}$. Además, $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$. Entonces

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3}$$

Por tanto, el conocimiento adicional de que el lanzamiento ha dado como resultado un número impar eleva la probabilidad de 1/2 a 2/3.

TEOREMAS DE PROBABILIDAD CONDICIONAL

Teorema 1-9 Dados tres eventos A_1, A_2, A_3, se tiene

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2) \tag{19}$$

Expresado en otras palabras, la probabilidad de que ocurran tanto A_1 como A_2 como A_3 es igual a la probabilidad que ocurra A_1 por la probabilidad de que ocurra A_2 dado que A_1 ha ocurrido, por la probabilidad de que ocurra A_3 dado que A_1 y A_2 han ocurrido. El resultado se generaliza fácilmente a n eventos.

Teorema 1-10 Si un suceso A debe ser el resultado de la ocurrencia de uno de los eventos mutuamente excluyentes $A_1, A_2, \ldots, A_n$, entonces

$$P(A) = P(A_1)\,P(A \mid A_1) + P(A_2)\,P(A \mid A_2) + \cdots + P(A_n)\,P(A \mid A_n) \tag{20}$$

EVENTOS INDEPENDIENTES

Si $P(B|A) = P(B)$, es decir, si la probabilidad de que B ocurra no se ve afectada por la ocurrencia o la no ocurrencia de A, entonces se dice que A y B son *eventos independientes*. Esto es equivalente a

$$P(A \cap B) = P(A)P(B) \tag{21}$$

de acuerdo con (18). Inversamente, si (21) se satisface, entonces A y B son independientes.

Se dice que tres eventos A_1, A_2, A_3, son *independientes* si son independientes por pares:

$$P(A_j \cap A_k) = P(A_j)P(A_k) \qquad j \neq k \quad \text{donde} \quad j, k = 1, 2, 3 \tag{22}$$

y

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) \tag{23}$$

Observar que ni (22) ni (23) son suficientes por sí solas. La independencia de más de tres eventos se define fácilmente.

TEOREMA O REGLA DE BAYES

Suponga que $A_1, A_2, \ldots, A_n$, son eventos mutuamente excluyentes cuya unión es el espacio muestral S, es decir, uno de estos sucesos debe ocurrir. Entonces, si A es un evento cualquiera, se tiene el importante teorema siguiente:

Teorema 1-11 **(regla de Bayes):**

$$P(A_k \mid A) = \frac{P(A_k)\,P(A \mid A_k)}{\sum_{j=1}^{n} P(A_j)\,P(A \mid A_j)} \tag{24}$$

Esto permite encontrar las probabilidades de los diversos eventos $A_1, A_2, \ldots, A_n$, que *hacen* que A ocurra. Por esta razón, al teorema de Bayes suele conocérsele como un *teorema sobre la probabilidad de causas*.

ANÁLISIS COMBINATORIO

En muchas ocasiones el número de los puntos muestrales en un espacio muestral no es muy grande, y la enumeración o la cuenta directa de los puntos muestrales, para obtener las probabilidades correspondientes es una tarea sencilla. Sin embargo, hay problemas en los que la cuenta directa resulta prácticamente imposible. En tales casos se hace uso del *análisis combinatorio*, al que también se le puede considerar como una *manera sofisticada de contar*.

PRINCIPIO FUNDAMENTAL DE CONTEO: DIAGRAMAS DE ÁRBOL

Si una acción puede realizarse de n_1 maneras diferentes y después una segunda acción puede realizarse de n_2 maneras distintas, . . . , y por último una k-ésima acción puede realizarse de n_k maneras diversas, entonces estas k acciones pueden realizarse, en el orden señalado, de $n_1 \cdot n_2 \cdots n_k$ maneras diferentes.

EJEMPLO 1.14 Si un hombre tiene 2 camisas y 4 corbatas, entonces tiene $2 \cdot 4 = 8$ maneras de elegir una camisa y una corbata.

Un diagrama, llamado por su aspecto *diagrama de árbol* (figura 1-4), suele emplearse como representación gráfica del principio fundamental de conteo, ya descrito.

EJEMPLO 1.15 Representando las camisas por S_1, S_2 y las corbatas por T_1, T_2, T_3, T_4, en el diagrama de árbol de la figura 1-4 se indican las diversas maneras de elegir una camisa y una corbata.

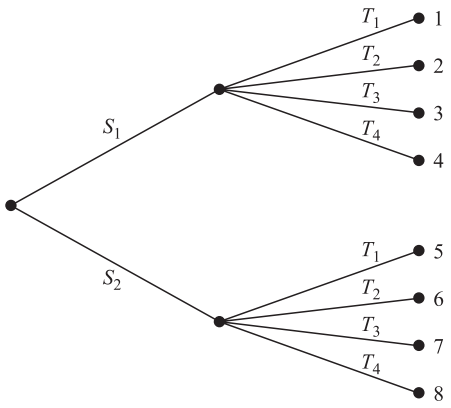


Figura 1-4

PERMUTACIONES

Suponga que se tienen n objetos diferentes y que se desea *ordenar* r de estos objetos uno tras otro en una línea. Como hay n maneras distintas de elegir el primer objeto y después $n - 1$ maneras diferentes de elegir el segundo objeto, . . . , y por último $n - r + 1$ maneras diversas de elegir el objeto r-ésimo, se deduce, de acuerdo con el principio fundamental de conteo, que la cantidad de *ordenamientos* diferentes o *permutaciones*, como suele llamárseles, está dada por

$$ {}_nP_r = n(n-1)(n-2)\cdots(n-r+1) \tag{25}$$

donde se observa que este producto tiene r factores. Al *número de permutaciones de n objetos tomados de r en r* se le denota ${}_nP_r$.

En el caso particular en el que $r = n$, (25) será

$$ {}_nP_n = n(n-1)(n-2)\cdots 1 = n! \tag{26}$$

a lo que se conoce como *factorial de n.* En términos de factoriales, (25) se escribe de la manera siguiente:

$$ {}_nP_r = \frac{n!}{(n-r)!} \tag{27}$$

Si $r = n$, se ve que (27) y (26) son iguales sólo si $0! = 1$, y esto se tomará como definición de 0!

EJEMPLO 1.16 El número de ordenamientos diferentes o permutaciones que consta de 3 letras tomados de las 7 letras: *A*, *B*, *C*, *D*, *E*, *F* y *G* es:

$$ {}_7P_3 = \frac{7!}{4!} = 7 \cdot 6 \cdot 5 = 210$$

Suponga que un conjunto consta de n objetos de los cuales n_1 son de un tipo (es decir, indistinguibles uno de otro), n_2 son de un segundo tipo, . . . , n_k son de un k-ésimo tipo. Aquí, por supuesto, $n = n_1 + n_2 + \cdots + n_k$. Entonces, el número de permutaciones diferentes de estos objetos es

$$ {}_nP_{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1!n_2!\cdots n_k!} \tag{28}$$

Vea el problema 1.25.

EJEMPLO 1.17 El número de permutaciones de 11 letras de la palabra *M I S S I S S I P P I*, la cual consta de una *M*, cuatro *I*, cuatro *S* y dos *P*, es:

$$ \frac{11!}{1!4!4!2!} = 34\,650$$

COMBINACIONES

En las permutaciones interesa el orden de los objetos. Por ejemplo, *abc* es una permutación diferente de *bca*. Sin embargo, en muchos problemas sólo estamos interesados en seleccionar o escoger objetos sin importar su orden. A estas elecciones se les llama *combinaciones*. Por ejemplo, *abc* y *bca* representan la misma combinación.

El número total de combinaciones de r objetos seleccionados de n objetos (que también se conoce como *combinaciones de n objetos tomados de r en r*) se denota ${}_nC_r$, o bien $\binom{n}{r}$. Se tiene (vea el problema 1.27)

$$ \binom{n}{r} = {}_nC_r = \frac{n!}{r!(n-r)!} \tag{29}$$

Lo que también puede escribirse como

$$ \binom{n}{r} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{{}_nP_r}{r!} \tag{30}$$

Es fácil demostrar que

$$ \binom{n}{r} = \binom{n}{n-r} \quad \text{o bien} \quad {}_nC_r = {}_nC_{n-r} \tag{31}$$

EJEMPLO 1.18 El número de maneras en las que se pueden tomar 3 cartas de 8 cartas diferentes es

$$ {}_8C_3 = \binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{3!} = 56 $$

COEFICIENTE BINOMIAL

A los números dados en (29) suele llamárseles *coeficientes binomiales* debido a que surgen en la *expansión binomial*

$$ (x + y)^n = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n}y^n \tag{32} $$

Estos coeficientes tienen muchas propiedades interesantes.

EJEMPLO 1.19

$$ \begin{aligned} (x + y)^4 &= x^4 + \binom{4}{1}x^3y + \binom{4}{2}x^2y^2 + \binom{4}{3}xy^3 + \binom{4}{4}y^4 \\ &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 \end{aligned} $$

APROXIMACIÓN DE STIRLING PARA $n!$

En los casos en que n es un número grande, no es fácil evaluar $n!$ En tales casos puede emplearse la fórmula de aproximación

$$ n! \sim \sqrt{2\pi n}\, n^n e^{-n} \tag{33} $$

en donde $e = 2.71828\ldots$, que es la base de los logaritmos naturales. El símbolo $\sim$ que aparece en (33) significa que el cociente del lado izquierdo entre lado derecho se aproxima a 1, a medida que $n \to \infty$.

El uso de las computadoras y las calculadoras ha opacado enormemente el valor de la fórmula de Stirling para cálculos numéricos, pero esta aproximación sigue siendo valiosa para estimaciones teóricas (vea el apéndice A).

PROBLEMAS RESUELTOS

EXPERIMENTOS ALEATORIOS, ESPACIOS MUESTRALES Y EVENTOS

1.1. De una baraja ordinaria de 52 cartas se saca al azar una carta. Describa el espacio muestral si los palos *a*) no se toman en cuenta, *b*) sí se toman en cuenta.

a) Si no se toman en cuenta los palos, el espacio muestral consta de as, dos, . . . , diez, sota, reina, rey, y puede indicarse $\{1, 2, \ldots, 13\}$.

b) Si se toman en cuenta los palos, el espacio muestral consta de as de corazones, de picas, de diamantes y de tréboles; . . . ; rey de corazones, de picas, de diamantes y de tréboles. Si se denotan los corazones, las picas, los diamantes y los tréboles, respectivamente, por 1, 2, 3, 4, por ejemplo, una sota de picas puede indicarse como (11, 2). El espacio muestral consta de los 52 puntos que se muestran en la figura 1-5.

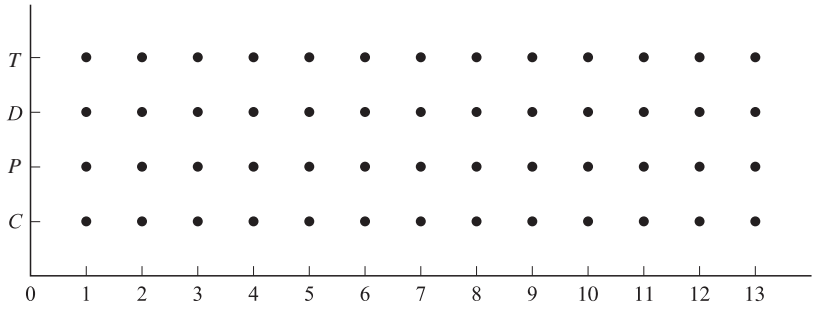


Figura 1-5

1.2. Volviendo al experimento del problema 1.1, sea A el evento {sacar un rey} o simplemente {rey} y B el evento {sacar un trébol} o simplemente {trébol}. Describir los eventos *a*) $A \cup B$, *b*) $A \cap B$, *c*) $A \cup B'$, *d*) $A' \cup B'$, *e*) $A - B$, *f*) $A' - B'$, *g*) $(A \cap B) \cup (A \cap B')$.

a) $A \cup B$ = {ya sea un rey o un trébol (o ambos, es decir, rey de tréboles)}.

b) $A \cap B$ = {un rey y trébol} = {rey de tréboles}.

c) Como B = {trébol}, B' = {no trébol} = {corazón, diamante, pica}.

Entonces $A \cup B'$ = {rey de corazones o de diamantes o de picas}.

d) $A' \cup B'$ = {no rey o no trébol} = {no rey de tréboles} = {cualquier carta que no sea rey de tréboles}.

Esto también lo vemos observando que $A' \cup B' = (A \cap B)'$ y usando *b*).

e) $A - B$ = {rey pero no trébol}.

Esto es lo mismo que $A \cap B'$ = {rey y no trébol}.

f) $A' - B'$ = {no rey y no "no trébol"} = {no rey y trébol} = {cualquier trébol excepto rey}.

Esto también lo vemos observando que $A' - B' = A' \cap (B')' = A' \cap B$.

g) $(A \cap B) \cup (A \cap B')$ = {(rey y trébol) o (rey y no trébol)} = {rey}.

Esto también lo vemos observando que $(A \cap B) \cup (A \cap B') = A$.

1.3. Utilizando la figura 1-5 describir los eventos *a*) $A \cup B$, *b*) $A' \cap B'$.

Los eventos buscados se indican en la figura 1-6. De manera similar, mediante estos diagramas también podemos indicar todos los eventos del problema 1.2. En la figura 1-6 observamos que $A' \cap B'$ es el complemento de $A \cup B$.

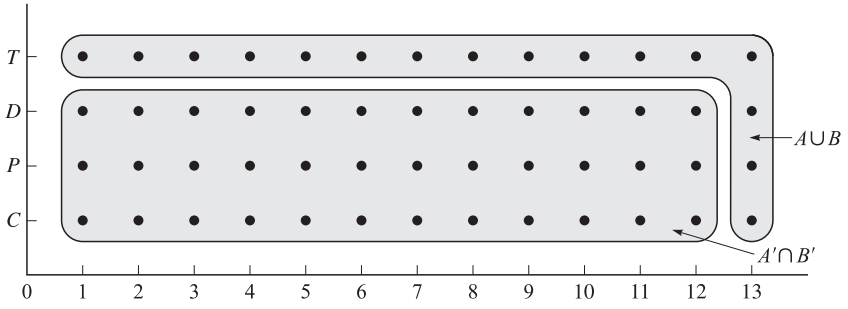


Figura 1-6

TEOREMAS SOBRE PROBABILIDAD

1.4. Demostrar *a*) el teorema 1-1, *b*) el teorema 1-2, *c*) el teorema 1-3, de las páginas 5 y 6.

a) Tenemos que $A_2 = A_1 \cup (A_2 - A_1)$ donde A_1 y $A_2 - A_1$ son mutuamente excluyentes. Entonces, de acuerdo con el axioma 3, página 5:

$$P(A_2) = P(A_1) + P(A_2 - A_1)$$

de manera que $$P(A_2 - A_1) = P(A_2) - P(A_1)$$

Como $P(A_2 - A_1) \geq 0$ por el axioma 1, página 5, también se concluye que $P(A_2) \geq P(A_1)$.

b) Por el axioma 1 sabemos que $P(A) \geq 0$. Para demostrar que $P(A) \leq 1$, observemos primero que $A \subseteq S$. Por tanto, por el teorema 1-1 [inciso *a*)] y por el axioma 2,

$$P(A) \leq P(S) = 1$$

c) Se tiene $S = S \cup \varnothing$. Como $S \cap \varnothing = \varnothing$, de acuerdo con el axioma 3 se deduce que

$$P(S) = P(S) + P(\varnothing) \qquad \text{o bien} \qquad P(\varnothing) = 0$$

1.5. Demostrar *a*) el teorema 1-4, *b*) el teorema 1-6.

a) Tenemos $A \cup A' = S$. Entonces como $A \cup A' = \varnothing$, tenemos

$$P(A \cup A') = P(S) \qquad \text{o bien} \qquad P(A) + P(A') = 1$$

es decir,

$$P(A') = 1 - P(A)$$

b) De acuerdo con el diagrama de Venn de la figura 1-7,

(1) $$A \cup B = A \cup [B - (A \cap B)]$$

Entonces, como los conjuntos A y $B - (A \cap B)$ son mutuamente excluyentes, tenemos, usando el axioma 3 y el teorema 1-1,

$$P(A \cup B) = P(A) + P[B - (A \cap B)]$$

$$= P(A) + P(B) - P(A \cap B)$$

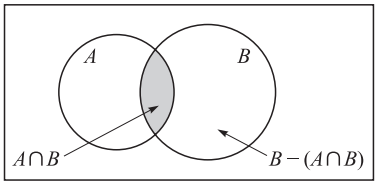


Figura 1-7

CÁLCULO DE PROBABILIDADES

1.6. De una baraja ordinaria de 52 cartas se extrae al azar una carta. Encontrar la probabilidad de que la carta sea *a*) un as, *b*) una sota de corazones, *c*) un 3 de tréboles o un 6 de diamantes, *d*) un corazón, *e*) cualquier palo, excepto corazones, *f*) un diez o una pica, *g*) ni un 4 ni un trébol.

Para simplificar se usarán C, P, D, T, para indicar corazones, picas, diamantes y tréboles, respectivamente, y 1, 2, . . . , 13 para as, 2, . . . , rey. Entonces $3 \cap C$ significará 3 de corazones, mientras que $3 \cup C$ significará tres o corazón. Usaremos el espacio muestral del problema 1.1*b*), asignando a cada punto muestral la misma probabilidad de $1/52$. Por ejemplo, $P(6 \cap T) = 1/52$.

a)
$$P(1) = P(1 \cap C \text{ o } 1 \cap P \text{ o } 1 \cap D \text{ o } 1 \cap T)$$

$$= P(1 \cap C) + P(1 \cap P) + P(1 \cap D) + P(1 \cap T)$$

$$= \frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{1}{13}$$

A esto también llegamos a partir del espacio muestral del problema 1.1*a*), en donde cada punto muestral, en particular el as, tiene probabilidad de $1/13$. Llegamos a lo mismo por un razonamiento sencillo de observar que hay 13 números y, por tanto, cada uno tiene una probabilidad de $1/13$ de ser extraído.

b) $P(11 \cap C) = \dfrac{1}{52}$

c) $P(3 \cap T \text{ o } 6 \cap D) = P(3 \cap T) + P(6 \cap D) = \dfrac{1}{52} + \dfrac{1}{52} = \dfrac{1}{26}$

d) $P(C) = P(1 \cap C \text{ o } 2 \cap C \text{ o } \ldots 13 \cap C) = \dfrac{1}{52} + \dfrac{1}{52} + \cdots + \dfrac{1}{52} = \dfrac{13}{52} = \dfrac{1}{4}$

A esto también se habría podido llegar observando que hay cuatro palos y que cada uno tiene la misma probabilidad de $1/4$ de ser extraído.

e) $P(C') = 1 - P(C) = 1 - \dfrac{1}{4} = \dfrac{3}{4}$ usando el inciso *d*) del teorema 1-4, página 6.

f) Como 10 y P no son mutuamente excluyentes, se tiene, de acuerdo con el teorema 1-6,

$$P(10 \cup P) = P(10) + P(P) - P(10 \cap P) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$$

g) La probabilidad de que ni 4 ni trébol se puede denotar como $P(4' \cap T')$. Pero $4' \cap T' = (4 \cup T)'$.

Por tanto,

$$P(4' \cap T') = P[(4 \cup T)'] = 1 - P(4 \cup T)$$
$$= 1 - [P(4) + P(T) - P(4 \cap T)]$$
$$= 1 - \left[\frac{1}{13} + \frac{1}{4} - \frac{1}{52}\right] = \frac{9}{13}$$

Esto también lo obtenemos observando que el diagrama favorable a este evento es el complemento del evento que se muestra marcado en la figura 1-8. Como este complemento tiene 52 – 16 = 36 puntos muestrales, a cada punto muestral se le asigna una probabilidad de 1/52, la probabilidad buscada es 36/52 = 9/13.

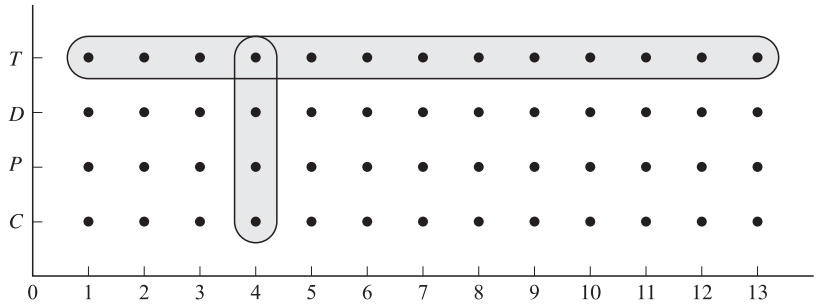


Figura 1-8

1.7. De una caja que contiene 6 bolas rojas, 4 blancas y 5 azules, se extrae al azar una bola. Determinar la probabilidad de que sea *a*) roja, *b*) blanca, *c*) azul, *d*) no roja, *e*) no blanca.

a) **Método 1**
Sean R, B y A los eventos de extraer una bola roja, una blanca y una azul, respectivamente. Entonces,

$$P(R) = \frac{\text{maneras de elegir una bola roja}}{\text{total de maneras de elegir una bola}} = \frac{6}{6 + 4 + 5} = \frac{6}{15} = \frac{2}{5}$$

Método 2
Este espacio muestral consta de $6 + 4 + 5 = 15$ puntos muestrales. Entonces, si asignamos a cada punto muestral la misma probabilidad de 1/15, vemos que $P(R) = 6/15 = 2/5$, dado que hay 6 puntos muestrales que corresponden a "bola roja".

b) $P(B) = \dfrac{4}{6 + 4 + 5} = \dfrac{4}{15}$

c) $P(A) = \dfrac{5}{6 + 4 + 5} = \dfrac{5}{15} = \dfrac{1}{3}$

d) $P(\text{no roja}) = P(R') = 1 - P(R) = 1 - \dfrac{2}{5} = \dfrac{3}{5}$ de acuerdo con el inciso *a*).

e) **Método 1**

$$P(\text{roja o blanca}) = P(R \cup B) = \frac{\text{maneras de elegir una bola roja o blanca}}{\text{total de maneras de elegir una bola}}$$
$$= \frac{6 + 4}{6 + 4 + 5} = \frac{10}{15} = \frac{2}{3}$$

Esto también se puede hacer empleando el espacio muestral del inciso *a*).

Método 2

$$P(R \cup B) = P(A') = 1 - P(A) = 1 - \frac{1}{3} = \frac{2}{3} \text{ de acuerdo con el inciso } c).$$

Método 3
Dado que los eventos R y B son mutuamente excluyentes, se deduce, de acuerdo con (4), página 5, que

$$P(R \cup B) = P(R) + P(B) = \frac{2}{5} + \frac{4}{15} = \frac{2}{3}$$

PROBABILIDAD CONDICIONAL Y EVENTOS INDEPENDIENTES

1.8. Un dado no cargado se lanza dos veces. Encontrar la probabilidad de obtener 4, 5 o 6 en el primer lanzamiento y 1, 2, 3 o 4 en el segundo.

Sea A_1 el evento "4, 5 o 6 en el primer lanzamiento", y sea A_2 el evento "1, 2, 3 o 4 en el segundo lanzamiento". Entonces se busca la probabilidad $P(A_1 \cap A_2)$.

Método 1

$$P(A_1 \cap A_2) = P(A_1)P(A_2|A_1) = P(A_1)P(A_2) = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = \frac{1}{3}$$

Aquí empleamos el hecho de que el resultado del segundo lanzamiento es *independiente* del resultado del primer lanzamiento, de manera que $P(A_2|A_1) = P(A_2)$. También usamos $P(A_1) = 3/6$ (ya que 4, 5 o 6 son 3 de 6 posibilidades igualmente probables) y $P(A_2) = 4/6$ (ya que 1, 2, 3 o 4 son 4 de 6 posibilidades igualmente probables).

Método 2

Cada una de las seis maneras en las que puede caer el dado en el primer lanzamiento puede combinarse con cada una de las seis maneras en las que puede caer en el segundo lanzamiento, en total $6 \cdot 6 = 36$ maneras, todas igualmente posibles.

Cada una de las tres maneras en las que puede ocurrir A_1 puede asociarse con cada una de las cuatro maneras en las que puede ocurrir A_2, dando en total $3 \cdot 4 = 12$ maneras en las que pueden ocurrir tanto A_1 como A_2. Entonces,

$$P(A_1 \cap A_2) = \frac{12}{36} = \frac{1}{3}$$

Esto muestra directamente que A_1 y A_2 son independientes ya que

$$P(A_1 \cap A_2) = \frac{1}{3} = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = P(A_1)P(A_2)$$

1.9. Encontrar la probabilidad de no obtener ni un 7 ni un 11 en ninguno de los dos lanzamientos de un par de dados no cargados.

En la figura 1-9 se muestra el espacio muestral correspondiente a cada lanzamiento de los dados. Por ejemplo, (5, 2) significa que con el primer dado se obtiene 5 y con el segundo se obtiene 2. Como los dados no están cargados hay 36 puntos muestrales, a cada uno de los cuales se le asigna la probabilidad de $1/36$.

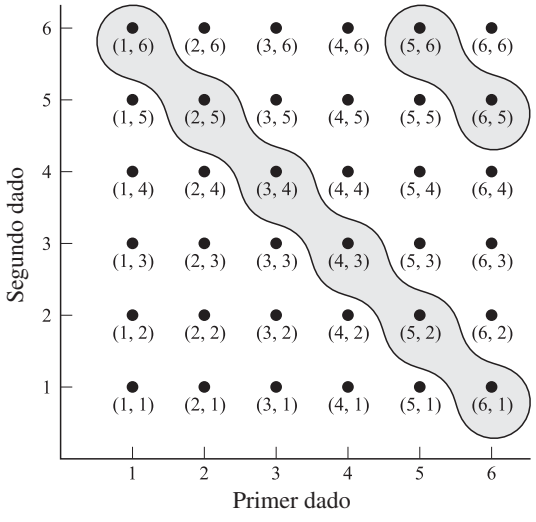


Figura 1-9

Si A es el evento "7 u 11", entonces A corresponde a los eventos encerrados en la figura 1-9. Como son 8 los puntos encerrados, tenemos $P(A) = 8/36 = 2/9$. Se concluye que la probabilidad de no tener ni un 7 ni un 11 está dada por

$$P(A') = 1 - P(A) = 1 - \frac{2}{9} = \frac{7}{9}$$

Empleando los subíndices 1 y 2 para denotar el primero y el segundo lanzamiento del dado, vemos que la probabilidad de no 7 u 11 ni en el primero ni en el segundo lanzamiento está dada por

$$P(A_1')\,P(A_2' \mid A_1') = P(A_1')P(A_2') = \left(\frac{7}{9}\right)\left(\frac{7}{9}\right) = \frac{49}{81},$$

con base en el hecho de que los lanzamientos son independientes.

1.10. De una baraja ordinaria bien mezclada de 52 cartas se extraen dos cartas. Encontrar la probabilidad de que las dos sean ases si la primera carta *a*) se repone, *b*) no se repone.

Método 1

Sea A_1 = evento "as en la primera extracción" y A_2 = evento "as en la segunda extracción". Entonces lo que se busca es $P(A_1 \cap A_2) = P(A_1)\,P(A_2|A_1)$.

a) Dado que para la primera extracción hay 4 ases en las 52 cartas, $P(A_1) = 4/52$. También, si la carta se repone antes de la segunda extracción, entonces $P(A_2|A_1) = 4/52$, ya que para la segunda extracción sigue habiendo cuatro ases en las 52 cartas. Entonces,

$$P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1) = \left(\frac{4}{52}\right)\left(\frac{4}{52}\right) = \frac{1}{169}$$

b) Como en el inciso *a*), $P(A_1) = 4/52$. Sin embargo, si en la primera extracción se obtiene un as, en las 51 cartas restantes habrá únicamente 3 ases, por tanto, $P(A_2|A_1) = 3/51$. Entonces

$$P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1) = \left(\frac{4}{52}\right)\left(\frac{3}{51}\right) = \frac{1}{221}$$

Método 2

a) La primera carta puede extraerse de 52 maneras distintas, y dado que hay reposición, la segunda carta también puede extraerse de 52 maneras distintas. Entonces, las dos cartas pueden extraerse de (52)(52) maneras distintas, todas igualmente posibles.

En tal caso hay 4 maneras de obtener un as en la primera extracción y 4 de obtener un as en la segunda extracción así que el número de maneras de obtener ases en la primera y en la segunda extracción es (4)(4). Entonces, la probabilidad buscada es

$$\frac{(4)(4)}{(52)(52)} = \frac{1}{169}$$

b) La primera carta puede extraerse de 52 maneras distintas, y dado que no hay reposición, la segunda carta puede extraerse de 51 formas distintas. Entonces, las dos cartas pueden extraerse de (52)(51) maneras distintas, todas igualmente posibles.

En tal caso hay 4 maneras de obtener un as en la primera extracción y 3 de obtener un as en la segunda extracción así que el número de maneras de obtener ases en la primera y en la segunda extracción es (4)(3). Entonces, la probabilidad buscada es

$$\frac{(4)(3)}{(52)(51)} = \frac{1}{221}$$

1.11. De la caja del problema 1.7 se extraen sucesivamente tres bolas. Encontrar la probabilidad de que se extraigan en el orden roja, blanca y azul si cada bola *a*) se repone, *b*) no se repone.

Sea R_1 = el evento "roja en la primera extracción", B_2 = el evento "blanca en la segunda extracción" y A_3 = el evento "azul en la tercera extracción". Buscamos $P(R_1 \cap B_2 \cap A_3)$.

a) Si cada bola se repone, entonces los eventos son independientes y

$$\begin{aligned} P(R_1 \cap B_2 \cap A_3) &= P(R_1)P(B_2 \mid R_1)P(A_3 \mid R_2 \cap B_2) \\ &= P(R_1)P(B_2)P(A_3) \\ &= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{6+4+5}\right)\left(\frac{5}{6+4+5}\right) = \frac{8}{225} \end{aligned}$$

b) Si no se repone cada bola, entonces los eventos son dependientes y

$$P(R_1 \cap B_2 \cap A_3) = P(R_1)P(B_2 \mid R_1)P(A_3 \mid R_1 \cap B_2)$$

$$= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{5+4+5}\right)\left(\frac{5}{5+3+5}\right) = \frac{4}{91}$$

1.12. Encontrar la probabilidad de obtener un 4 en dos lanzamientos de un dado legal.

Sea A_1 = evento "4 en el primer lanzamiento" y A_2 = evento "4 en el segundo lanzamiento". Entonces

$A_1 \cup A_2$ = evento "4 en el primer lanzamiento o 4 en el segundo lanzamiento o en ambos"
= evento "por lo menos un 4"

y lo que se busca es $P(A_1 \cup A_2)$.

Método 1

Los eventos A_1 y A_2 no son mutuamente excluyentes pero son independientes. Por tanto, de acuerdo con (10) y (21),

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

$$= P(A_1) + P(A_2) - P(A_1)P(A_2)$$

$$= \frac{1}{6} + \frac{1}{6} - \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{11}{36}$$

Método 2

$$P(\text{se obtiene por lo menos un 4}) + P(\text{no se obtiene ningún 4}) = 1$$

Entonces $P(\text{se obtiene por lo menos un 4}) = 1 - P(\text{no se obtiene ningún 4})$

$= 1 - P(\text{no se obtiene 4 en el primer lanzamiento y no se obtiene 4 en el segundo lanzamiento})$

$$= 1 - P(A_1' \cap A_2') = 1 - P(A_1')P(A_2')$$

$$= 1 - \left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{11}{36}$$

Método 3

El número total de maneras igualmente factibles en las que pueden caer los dos dados es $= 6 \cdot 6 = 36$.

Además:
Número de maneras en las que puede ocurrir A_1 pero no $A_2 = 5$
Número de maneras en las que puede ocurrir A_2 pero no $A_1 = 5$
Número de maneras en las que pueden ocurrir A_1 y $A_2 = 1$

Por tanto, el número de maneras en las que puede ocurrir por lo menos uno de los eventos A_1 o $A_2 = 5 + 5 + 1 = 11$. Por tanto, $P(A_1 \cup A_2) = 11/36$.

1.13. Una bolsa contiene 4 bolas blancas y 2 bolas negras; otra contiene 3 bolas blancas y 5 bolas negras. Si de cada bolsa se extrae una bola, encontrar la probabilidad de que *a*) las dos sean blancas, *b*) las dos sean negras, *c*) una sea blanca y la otra sea negra.

Sean W_1 = evento "bola blanca de la primera bolsa", W_2 = evento "bola blanca de la segunda bolsa ".

a) $P(W_1 \cap W_2) = P(W_1)P(W_2 \mid W_1) = P(W_1)P(W_2) = \left(\frac{4}{4+2}\right)\left(\frac{3}{3+5}\right) = \frac{1}{4}$

b) $P(W_1' \cap W_2') = P(W_1')P(W_2' \mid W_1') = P(W_1')P(W_2') = \left(\frac{2}{4+2}\right)\left(\frac{5}{3+5}\right) = \frac{5}{24}$

c) La probabilidad buscada es

$$1 - P(W_1 \cap W_2) - P(W_1' \cap W_2') = 1 - \frac{1}{4} - \frac{5}{24} = \frac{13}{24}$$

1.14. Demostrar el teorema 1-10, página 7.

El teorema se demuestra para el caso $n = 2$. Extensiones para n mayores son sencillas. Si un evento A resulta en dos eventos mutuamente excluyentes A_1, A_2, entonces

$$A = (A \cap A_1) \cup (A \cap A_2)$$

Pero $A \cap A_1$ y $A \cap A_2$ son mutuamente excluyentes ya que A_1 y A_2 lo son. Por tanto, de acuerdo con el axioma 3,

$$P(A) = P(A \cap A_1) + P(A \cap A_2)$$
$$= P(A_1)\,P(A \mid A_1) + P(A_2)\,P(A \mid A_2)$$

empleando (18), página 7.

1.15. La caja *I* contiene 3 canicas rojas y 2 azules. La caja *II* contiene 2 canicas rojas y 8 azules. Se lanza una moneda no cargada. Si cae cara se toma una canica de la caja *I*; si cae cruz, se toma una canica de la caja *II*. Encuentre la probabilidad de tomar una canica roja.

Sea *R* el evento "se toma una canica roja", y *I* y *II* sean los eventos se toma de la caja *I* o de la caja *II*, respectivamente. Dado que una canica roja puede tomarse de la caja *I* o de la caja *II*, pueden emplearse los resultados del problema 1.14 con $A = R, A_1 = I, A_2 = II$. Por tanto, la probabilidad de tomar una canica roja es

$$P(R) = P(I)P(R \mid I) + P(II)P(R \mid II) = \left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{2+8}\right) = \frac{2}{5}$$

TEOREMA DE BAYES

1.16. Demostrar el teorema de Bayes (teorema 1-11, página 8).

Como *A* resulta en uno de los eventos mutuamente excluyentes $A_1, A_2, \ldots, A_n$, se tiene, de acuerdo con el teorema 1-10 (problema 1.14),

$$P(A) = P(A_1)P(A \mid A_1) + \cdots + P(A_n)P(A \mid A_n) = \sum_{j=1}^{n} P(A_j)P(A \mid A_j)$$

Por tanto,
$$P(A_k \mid A) = \frac{P(A_k \cap A)}{P(A)} = \frac{P(A_k)P(A \mid A_k)}{\sum_{j=1}^{n} P(A_j)P(A \mid A_j)}$$

1.17. En el problema 1.15, se supone que la persona que lanza la moneda no dice si cayó cara o cruz (de manera que no se sabe de qué caja se tomó la canica) pero sí dice que se tomó una canica roja. ¿Cuál es la probabilidad de que se haya tomado de la caja *I* (es decir, que la moneda cayó cara)?

Usaremos la misma notación que en el problema 1.15, es decir, $A = R, A_1 = I, A_2 = II$. Buscamos la probabilidad de que se haya sacado de la caja *I* dado que se sabe que la canica tomada es roja. Empleando la regla de Bayes con $n = 2$, está probabilidad está dada por

$$P(I \mid R) = \frac{P(I)P(R \mid I)}{P(I)P(R \mid I) + P(II)P(R \mid II)} = \frac{\left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right)}{\left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{2+8}\right)} = \frac{3}{4}$$

ANÁLISIS COMBINATORIO, CONTEO Y DIAGRAMAS DE ÁRBOL

1.18. Se va a formar un comité de 3 personas que constará de un representante de los trabajadores, uno de los administradores y uno del público. Si hay 3 posibles representantes de los trabajadores, 2 de los administradores y 4 del público, determine cuántos son los comités que pueden formarse usando *a*) el principio fundamental de conteo y *b*) un diagrama de árbol.

a) Al representante de los trabajadores se le puede escoger de 3 maneras diferentes y después de esto al representante de los administradores se le puede escoger de 2 modos distintos. Entonces, hay $3 \cdot 2 = 6$ maneras diferentes de escoger un representante de los trabajadores y uno de los administradores. Dada cualquiera de estas maneras, se puede escoger un representante del público de cuatro formas diferentes. Por tanto, el número de comités distintos que pueden formarse es: $3 \cdot 2 \cdot 4 = 24$.

b) Se denota a los 3 representantes de los trabajadores como L_1, L_2, L_3; a los representantes de los administradores como M_1, M_2, y a los representantes del público como P_1, P_2, P_3, P_4. Entonces, el diagrama de árbol de la figura 1-10 muestra que en total hay 24 comités. A partir de este diagrama se pueden enumerar todos los comités posibles, es decir, $L_1M_1P_1$, $L_1M_1P_2$, etcétera.

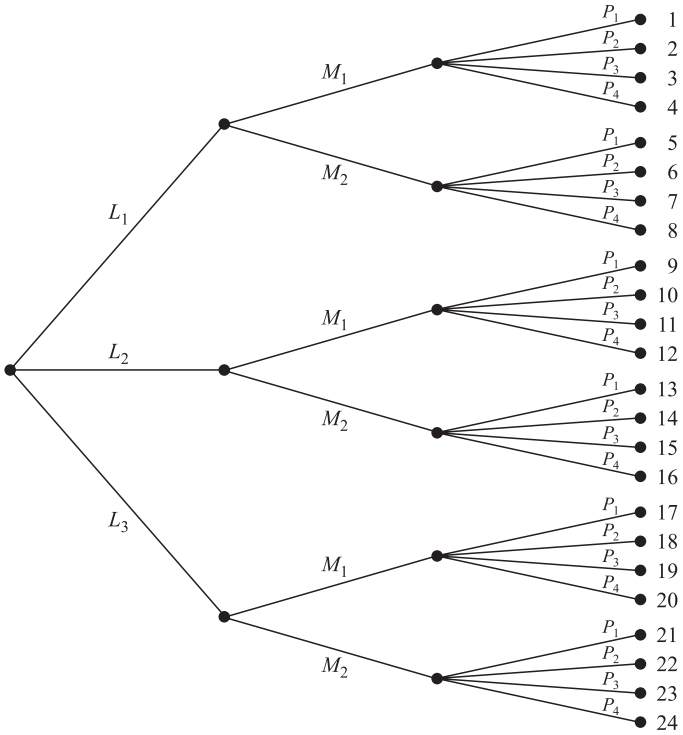


Figura 1-10

PERMUTACIONES

1.19. ¿De cuántas maneras pueden ponerse en hilera cinco canicas diferentes?

Hay que ordenar las 5 canicas en cinco posiciones así: $-\,-\,-\,-\,-$. La primera posición puede ocuparla cualquiera de las 5 canicas, es decir, hay cinco maneras de ocupar la primera posición. Una vez hecho esto, hay 4 maneras de ocupar la segunda posición. Después hay 3 maneras de ocupar la tercera posición, 2 de ocupar la cuarta posición y, por último, sólo 1 manera de ocupar la última posición. Por tanto:

$$\text{Número de maneras en que se pueden colocar 5 canicas en hilera} = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$$

En general,

$$\text{número de maneras en que se pueden colocar } n \text{ objetos en hilera} = n(n-1)(n-2)\cdots 1 = n!$$

A esto también se le conoce como *número de permutaciones de n objetos diferentes tomados de n en n* y se denota como ${}_nP_n$.

1.20. ¿De cuántas maneras pueden sentarse 10 personas en un banco en el que sólo hay asiento para 4?

El primer asiento puede ocuparse de 10 maneras, y una vez hecho esto, hay 9 para ocupar el segundo asiento, 8 maneras para ocupar el tercer asiento y 7 para ocupar el cuarto asiento. Por tanto,

$$\text{Número de maneras en que se pueden ordenar 10 personas tomadas de 4 en 4} = 10 \cdot 9 \cdot 8 \cdot 7 = 5\,040$$

En general,

$$\text{Número de maneras en que se pueden ordenar } n \text{ objetos diferentes tomados de } r \text{ en } r = n(n-1)\cdots(n-r+1).$$

A esto también se le conoce como *número de permutaciones de n objetos tomados de r en r* y se denota como ${}_nP_r$. Observe que cuando $r = n$, ${}_nP_n = n!$, como en el problema 1.19.

1.21. Evaluar *a*) ${}_8P_3$, *b*) ${}_6P_4$, *c*) ${}_{15}P_1$, *d*) ${}_3P_3$.

a) ${}_8P_3 = 8 \cdot 7 \cdot 6 = 336$ *b*) ${}_6P_4 = 6 \cdot 5 \cdot 4 \cdot 3 = 360$ *c*) ${}_{15}P_1 = 15$ *d*) ${}_3P_3 = 3 \cdot 2 \cdot 1 = 6$

1.22. Se desea sentar a 5 hombres y a 4 mujeres en una fila, uno junto a otro, de modo que las mujeres ocupen los lugares pares. ¿De cuántas maneras se pueden sentar?

Los hombres se pueden acomodar de ${}_5P_5$ maneras y las mujeres de ${}_4P_4$ maneras. Cada forma de sentar a los hombres puede asociarse con cada manera de sentar a las mujeres. Por tanto,

$$\text{Maneras de sentarlos} = {}_5P_5 \cdot {}_4P_4 = 5!\,4! = (120)(24) = 2\,880$$

1.23. ¿Cuántos números de 4 dígitos se pueden formar con los 10 dígitos 0, 1, 2, 3, … , 9 si *a*) se permiten las repeticiones, *b*) no se permiten las repeticiones, *c*) el último dígito debe ser cero y no se permiten las repeticiones?

a) El primer dígito puede ser cualquiera de 9 dígitos (puesto que no se permite el 0). El segundo, tercero y cuarto dígitos pueden ser cualquiera de los 10 dígitos. Entonces podemos formar $9 \cdot 10 \cdot 10 \cdot 10 = 9\,000$ números.

b) El primer dígito puede ser cualquiera del 1 al 9 (cualquiera menos el 0).

El segundo dígito puede ser cualquiera del 1 al 9 (cualesquiera, excepto el usado como primer dígito).

El tercer dígito puede ser cualquiera de 8 (cualesquiera menos los usados para los primeros dos dígitos).

El cuarto dígito puede ser cualquiera de 7 dígitos (cualesquiera salvo los usados para los primeros tres dígitos).

Entonces $9 \cdot 9 \cdot 8 \cdot 7 = 4\,536$, números que se pueden formar.

Otro método
El primer dígito puede ser cualquiera del 1 al 9, y los tres restantes los podemos elegir de ${}_9P_3$ maneras. Entonces, $9 \cdot {}_9P_3 = 9 \cdot 9 \cdot 8 \cdot 7 = 4\,536$ números que podemos formar.

c) El primer dígito lo podemos elegir de 9 maneras, el segundo de 8, y el tercero de 7 maneras. Entonces, $9 \cdot 8 \cdot 7 = 504$, números que podemos formar.

Otro método
El primer dígito lo podemos elegir de 9 maneras, y los dos dígitos siguientes de ${}_8P_2$ maneras. Entonces, $9 \cdot {}_8P_2 = 9 \cdot 8 \cdot 7 = 504$, números que podemos formar.

1.24. Se tienen cuatro libros distintos de matemáticas, seis libros diferentes de física, y dos libros distintos de química que se deben acomodar en un estante. ¿De cuántas maneras pueden acomodarse si *a*) los libros de cada tema deben estar juntos, *b*) solamente los libros de matemáticas deben estar juntos?

a) Los libros de matemáticas los podemos ordenar entre sí de ${}_4P_4 = 4!$ maneras, los libros de física de ${}_6P_6 = 6!$, los de química de ${}_2P_2 = 2!$ y los tres grupos de ${}_3P_3 = 3!$ maneras. Por tanto,

$$\text{Número de maneras en las que podemos acomodarlos} = 4!6!2!3! = 207\,360.$$

b) Consideremos los cuatro libros de matemáticas como un solo libro grande. Entonces tenemos 9 libros, los cuales podemos acomodar de ${}_9P_9 = 9!$ maneras. En todas éstas los libros de matemáticas están juntos. Pero los libros de matemáticas, entre sí, los podemos acomodar de ${}_4P_4 = 4!$ maneras. Por tanto,

$$\text{Número de maneras en las que podemos acomodar los libros} = 9!4! = 8\,709\,120$$

1.25. Hay cinco canicas rojas, dos blancas y tres azules en una hilera. Si todas las canicas del mismo color no se pueden distinguir una de otra, ¿de cuántas maneras distintas pueden colocarse en hilera las canicas?

Supongamos que hay N maneras diferentes de acomodar las canicas. Multiplicando N por los números de maneras de acomodar *a*) las cinco rojas entre sí, *b*) las dos blancas entre sí y *c*) las tres azules entre sí (es decir, multiplicando N por 5!2!3!), obtenemos el número de maneras en que podemos colocar en hilera las 10 canicas si fueran todas distinguibles, es decir, 10!

Entonces, $(5!2!3!)N = 10!$ y $N = 10!/(5!2!3!)$

En general la cantidad de maneras diferentes en que se pueden ordenar n objetos de los cuales n_1 son iguales, n_2 son iguales,… , n_k son iguales es $\dfrac{n!}{n_1!n_2! \cdots n_k!}$ donde $n_1 + n_2 + \cdots + n_k = n$.

1.26. ¿De cuántas maneras se pueden sentar 7 personas en una mesa redonda si *a*) pueden sentarse como quieran, *b*) 2 determinadas personas no deben sentarse juntas?

a) Si se deja que uno de ellos se siente donde quiera. Entonces las 6 personas restantes se pueden acomodar de $6! = 720$ maneras, que es el número total de formas de acomodar a las 7 personas en círculo.

b) Considere a 2 determinadas personas como una persona. Entonces hay 6 personas en total que pueden ser acomodadas de 5! maneras. Pero las 2 personas consideradas como una se les pueden acomodar de 2! maneras. Por tanto, el número de maneras en que se pueden acomodar 7 personas en una mesa redonda sentando juntas a 2 determinadas personas $= 5!2! = 240$.

Después usando *a*), el número total de maneras en las cuales 7 personas se pueden sentar en una mesa redonda de modo que 2 determinadas personas no se sienten juntas $= 730 - 240 = 490$ maneras.

COMBINACIONES

1.27. ¿De cuántas maneras se pueden dividir 10 objetos en dos grupos que contengan 4 y 6 objetos, respectivamente?

Esto es lo mismo que el número de arreglos de 10 objetos de los cuales 4 son semejantes y los otros 6 son objetos semejantes. De acuerdo con el problema 1.25, esto es $\dfrac{10!}{4!6!} = \dfrac{10 \cdot 9 \cdot 8 \cdot 7}{4!} = 210$.

Este problema es equivalente a encontrar el número de maneras de escoger 4 de 10 objetos (o 6 de 10 objetos), siendo inmaterial el orden en el que se elijan. En general, al número de maneras de elegir r de n objetos, se le llama *número de combinaciones de n objetos tomados r a la vez* y se denota ${}_nC_r$ o $\binom{n}{r}$, y está dado por

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{{}_nP_r}{r!} $$

1.28. Evaluar *a*) ${}_7C_4$, *b*) ${}_6C_5$, *c*) ${}_4C_4$.

a) $ {}_7C_4 = \dfrac{7!}{4!3!} = \dfrac{7 \cdot 6 \cdot 5 \cdot 4}{4!} = \dfrac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 35. $

b) $ {}_6C_5 = \dfrac{6!}{5!1!} = \dfrac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2}{5!} = 6, \quad \text{o} \quad {}_6C_5 = {}_6C_1 = 6. $

c) ${}_4C_4$ es el número de maneras de elegir 4 objetos tomando 4 a la vez, y sólo hay una manera. Entonces ${}_4C_4 = 1$. Observemos que formalmente

$$ {}_4C_4 = \frac{4!}{4!0!} = 1 \quad \text{si se } \textit{define } 0! = 1. $$

1.29. ¿De cuántas maneras se puede elegir un comité de 5 personas a partir de 9 personas?

$$ \binom{9}{5} = {}_9C_5 = \frac{9!}{5!4!} = \frac{9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}{5!} = 126 $$

1.30. A partir de 5 matemáticos y 7 físicos, se debe formar un comité que conste de 2 matemáticos y 3 físicos. ¿De cuántas maneras puede hacerse esto si *a*) cualquier matemático y cualquier físico pueden incluirse, *b*) un físico particular debe estar en el comité, *c*) dos matemáticos en particular no pueden estar en el comité?

a) 2 de 5 matemáticos se pueden seleccionar de ${}_5C_2$ maneras.
3 de 7 físicos se pueden seleccionar de ${}_7C_3$ maneras.

$$ \text{Número total de selecciones posibles} = {}_5C_2 \cdot {}_7C_3 = 10 \cdot 35 = 350 $$

b) 2 de 5 matemáticos se pueden seleccionar de ${}_5C_2$ maneras.
2 de 6 físicos se pueden seleccionar de ${}_6C_2$ maneras.

$$ \text{Número total de selecciones posibles} = {}_5C_2 \cdot {}_6C_2 = 10 \cdot 15 = 150 $$

c) 2 de 3 matemáticos se pueden seleccionar de ${}_3C_2$ maneras.
3 de 7 físicos se pueden seleccionar de ${}_7C_3$ maneras.

$$ \text{Número total de selecciones posibles} = {}_3C_2 \cdot {}_7C_3 = 3 \cdot 35 = 105 $$

1.31. ¿Cuántas ensaladas se pueden hacer con lechuga, escarola, endibia, berro y achicoria?

Cada ingrediente se puede tratar de 2 maneras, elegir tratarlo o no. Puesto que cada una de estas 2 maneras se asocia con 2 modos de emplear o no a cada uno de los otros ingredientes, el número de maneras de emplear los 5 ingredientes $= 2^5$ maneras éstas incluyen el caso en el que no se elige ninguno de los ingredientes. Por tanto,

$$\text{Número de ensaladas} = 2^5 - 1 = 31$$

Otro método
Uno puede seleccionar 5 ingredientes de uno cualquiera, 2 de 5 ingredientes, . . . , 5 de 5 ingredientes. Entonces el número de ensaladas que obtenemos es

$${}_5C_1 + {}_5C_2 + {}_5C_3 + {}_5C_4 + {}_5C_5 = 5 + 10 + 10 + 5 + 1 = 31$$

En general para cualquier número entero positivo n, ${}_nC_1 + {}_nC_2 + {}_nC_3 + \cdots + {}_nC_n = 2^n - 1$.

1.32. A partir de 7 consonantes y 5 vocales, ¿cuántas palabras se pueden formar que consten de 4 diferentes consonantes y 3 vocales distintas? Las palabras no necesitan tener significado.

Las 4 diferentes consonantes se pueden seleccionar de ${}_7C_4$ maneras, las 3 vocales distintas se pueden seleccionar de ${}_5C_3$ maneras, y las 7 diferentes letras que se obtienen (4 consonantes, 3 vocales) se pueden ordenar entre sí de ${}_7P_7 = 7!$ maneras. Entonces,

$$\text{Número de palabras} = {}_7C_4 \cdot {}_5C_3 \cdot 7! = 35 \cdot 10 \cdot 5\,040 = 1\,764\,000$$

COEFICIENTES BINOMIALES

1.33. Demostrar que $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$.

Se tiene

$$\begin{aligned}\binom{n}{r} = \frac{n!}{r!(n-r)!} &= \frac{n(n-1)!}{r!(n-r)!} = \frac{(n-r+r)(n-1)!}{r!(n-r)!} \\ &= \frac{(n-r)(n-1)!}{r!(n-r)!} + \frac{r(n-1)!}{r!(n-r)!} \\ &= \frac{(n-1)!}{r!(n-r-1)!} + \frac{(n-1)!}{(r-1)!(n-r)!} \\ &= \binom{n-1}{r} + \binom{n-1}{r-1}\end{aligned}$$

Este resultado tiene el siguiente uso interesante. Si se escriben los coeficientes del desarrollo binomial $(x + y)^n$ para $n = 0, 1, 2, \ldots$, se obtiene el arreglo siguiente, llamado *triángulo de Pascal*:

$$\begin{array}{ll}
n = 0 & \qquad\qquad\qquad 1 \\
n = 1 & \qquad\qquad 1 \quad 1 \\
n = 2 & \qquad\quad 1 \quad 2 \quad 1 \\
n = 3 & \qquad 1 \quad 3 \quad 3 \quad 1 \\
n = 4 & \quad 1 \quad 4 \quad 6 \quad 4 \quad 1 \\
n = 5 & \quad 1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
n = 6 & \quad 1 \quad 6 \quad 15 \quad 20 \quad 15 \quad 6 \quad 1 \\
\text{etc.} &
\end{array}$$

Cualquier número de cualquier renglón lo obtenemos sumando los dos números del renglón precedente, que estén más cerca a su izquierda y a su derecha. Así, $10 = 4 + 6$, $15 = 10 + 5$, etcétera.

1.34. Encontrar el término constante en la expansión $\left(x^2 + \frac{1}{x}\right)^{12}$.

De acuerdo con el teorema del binomio,

$$\left(x^2 + \frac{1}{x}\right)^{12} = \sum_{k=0}^{12}\binom{12}{k}(x^2)^k\left(\frac{1}{x}\right)^{12-k} = \sum_{k=0}^{12}\binom{12}{k}x^{3k-12}.$$

El término constante es aquel para el que $3k - 12 = 0$, esto es $k = 4$ y, por tanto, está dado por

$$\binom{12}{4} = \frac{12 \cdot 11 \cdot 10 \cdot 9}{4 \cdot 3 \cdot 2 \cdot 1} = 495$$

PROBABILIDAD USANDO ANÁLISIS COMBINATORIO

1.35. Una caja contiene 8 bolas rojas, 3 blancas y 9 azules. Si se extraen 3 bolas al azar sin reposición, determine la probabilidad de que *a*) las 3 sean rojas, *b*), las 3 sean blancas, *c*) 2 sean rojas y 1 sea blanca, *d*) por lo menos 1 sea blanca, *e*) sea 1 de cada color, *f*) las bolas se extraigan en el orden roja, blanca, azul.

a) **Método 1**
Sean R_1, R_2, R_3 los eventos, "bola roja en la primera extracción", "bola roja en la segunda extracción", "bola roja en la tercera extracción", respectivamente. Entonces $R_1 \cap R_2 \cap R_3$ denota el evento "las 3 bolas extraídas son rojas." Por tanto, se tiene

$$P(R_1 \cap R_2 \cap R_3) = P(R_1)P(R_2 \mid R_1)P(R_3 \mid R_1 \cap R_2)$$
$$= \left(\frac{8}{20}\right)\left(\frac{7}{19}\right)\left(\frac{6}{18}\right) = \frac{14}{285}$$

Método 2

$$\text{Probabilidad requerida} = \frac{\text{número de maneras de seleccionar 3 de 8 bolas rojas}}{\text{número de maneras de seleccionar 3 de 20 bolas}} = \frac{{}_8C_3}{{}_{20}C_3} = \frac{14}{285}$$

b) Usando el segundo método del inciso *a*),

$$P(\text{las 3 sean blancas}) = \frac{{}_3C_3}{{}_{20}C_3} = \frac{1}{1\,140}$$

También se puede usar el primer método del inciso *a*).

c) $P(\text{2 sean rojas y 1 sea blanca})$

$$= \frac{(\text{seleccionar 2 de 8 bolas rojas})(\text{selecciones 1 de 3 bolas blancas})}{\text{número de selecciones de 3 de 20 bolas}}$$
$$= \frac{({}_8C_2)({}_3C_1)}{{}_{20}C_3} = \frac{7}{95}$$

d) $P(\text{ninguna sea blanca}) = \frac{{}_{17}C_3}{{}_{20}C_3} = \frac{34}{57}$. Entonces,

$$P(\text{por lo menos 1 sea blanca}) = 1 - \frac{34}{57} = \frac{23}{57}$$

e) $P(\text{sea extraída una de cada color}) = \frac{({}_8C_1)({}_3C_1)({}_9C_1)}{{}_{20}C_3} = \frac{18}{95}$

f) $P(\text{las bolas se extraigan en la orden roja, blanca, azul}) = \frac{1}{3!}P(\text{una sea extraída de cada color})$

$$= \frac{1}{6}\left(\frac{18}{95}\right) = \frac{3}{95}, \text{ usando } e)$$

Otro método

$$P(R_1 \cap B_2 \cap A_3) = P(R_1)P(B_2 \mid R_1)P(A_3 \mid R_1 \cap B_2)$$
$$= \left(\frac{8}{20}\right)\left(\frac{3}{19}\right)\left(\frac{9}{18}\right) = \frac{3}{95}$$

1.36. En el juego de póquer se extraen 5 cartas de un paquete de 52 cartas bien mezcladas. Encontrar la probabilidad de *a*) 4 sean ases y cualquier otra carta, *b*) 4 sean ases y 1 sea rey, *c*) 3 sean dieces y 2 sean sotas, *d*) extraer nueve, diez, sota, reina, rey en cualquier orden, *e*) 3 sean de una figura cualquiera y 2 sean de otra, *f*) extraer por lo menos 1 as.

a) $P(\text{4 ases}) = \dfrac{({}_4C_4)({}_{48}C_1)}{{}_{52}C_5} = \dfrac{1}{54\,145}.$

b) $P(\text{4 ases y 1 rey}) = \dfrac{({}_4C_4)({}_4C_1)}{{}_{52}C_5} = \dfrac{1}{649\,740}.$

c) $P(\text{3 sean dieces y 2 sean sotas}) = \dfrac{({}_4C_3)({}_4C_2)}{{}_{52}C_5} = \dfrac{1}{108\,290}.$

d) $P(\text{nueve, diez, sota, reina, rey en cualquier orden}) = \dfrac{({}_4C_1)({}_4C_1)({}_4C_1)({}_4C_1)({}_4C_1)}{{}_{52}C_5} = \dfrac{64}{162\,435}.$

e) $P(\text{3 de una figura cualquiera, y 2 de otra}) = \dfrac{(4 \cdot {}_{13}C_3)(3 \cdot {}_{13}C_2)}{{}_{52}C_5} = \dfrac{429}{4\,165},$

puesto que hay 4 maneras de elegir la primera figura y 3 maneras de elegir la segunda figura.

f) $P(\text{ningún as}) = \dfrac{{}_{48}C_5}{{}_{52}C_5} = \dfrac{35\,673}{54\,145}.$ Entonces $P(\text{por lo menos un as}) = 1 - \dfrac{35\,673}{54\,145} = \dfrac{18\,472}{54\,145}.$

1.37. Determinar la probabilidad de obtener 3 seises en 5 lanzamientos con un dado no cargado.

Representar los lanzamientos del dado por los 5 espacios $-\,-\,-\,-\,-$. En cada espacio se tendrá el evento 6 o no 6 (6′). Por ejemplo, 3 seises y 2 no seises pueden presentarse de la forma 6 6 6′6 6′ o bien 6 6′ 6 6′ 6, etc. Ahora la probabilidad del resultado 6 6 6′ 6 6′ es

$$P(6\ 6\ 6'\ 6\ 6') = P(6)\,P(6)\,P(6')\,P(6)\,P(6') = \frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\cdot\frac{5}{6} = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2$$

puesto que se supone independencia. De manera similar

$$P = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2$$

para el resto de los resultados en los cuales se obtengan 3 seis y 2 no seises. Pero hay ${}_5C_3 = 10$ de estos resultados, y éstos son mutuamente excluyentes. Por tanto, la probabilidad requerida es

$$P(6\ 6\ 6'6\ 6'\ \text{o}\ \ 6\ 6'6\ 6'6\ \text{o}\ldots) = {}_5C_3\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = \frac{5!}{3!2!}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = \frac{125}{3\,888}$$

En general si $p = P(A)$ y $q = 1 - p = P(A')$, entonces usando el mismo razonamiento anterior, la probabilidad de obtener exactamente x cantidad de A en n ensayos independientes es

$${}_nC_x p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}$$

1.38. Un estante tiene 6 libros de matemáticas y 4 libros de física. Encuentre la probabilidad de que 3 libros particulares de matemáticas estén juntos.

Todos los libros se pueden ordenar entre sí de ${}_{10}P_{10} = 10!$ maneras. Suponga que los 3 libros particulares de matemáticas sean sustituidos por un libro. Entonces se tiene un total de 8 libros que se puedan ordenar de ${}_8P_8 = 8!$ maneras. Pero los 3 de matemáticas se pueden ordenar de ${}_3P_3 = 3!$ maneras. Por tanto, la probabilidad requerida está dada por

$$\frac{8!\,3!}{10!} = \frac{1}{15}$$

PROBLEMAS DIVERSOS

1.39. A y B juegan 12 partidas de ajedrez, de los cuales 6 los gana A, 4 los gana B, y 2 terminan en empate. Acuerdan jugar un torneo de 3 juegos. Encontrar la probabilidad de que *a*) A gane los 3 juegos, *b*) 2 juegos terminen en empate, *c*) A y B ganen alternativamente, *d*) B gane por lo menos un juego.

Sean A_1 A_2, A_3 los eventos "A gana" el primero, el segundo y el tercer juegos, respectivamente, y sean B_1, B_2, B_3 los eventos "B gana" el primero, el segundo y el tercer juegos, respectivamente. Con base en su último desempeño (probabilidad empírica),

supondremos que

$$P(A \text{ gana alguno de los juegos}) = \frac{6}{12} = \frac{1}{2}, \qquad P(B \text{ gana alguno de los juegos}) = \frac{4}{12} = \frac{1}{3}$$

a) $P(A \text{ gana los 3 juegos}) = P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2)\,P(A_3) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}$

si se supone que los resultados de cada juego son independientes de los resultados de cualquiera de los otros juegos. (Esta suposición no estará justificada si algún jugador se ve *influenciado psicológicamente* porque el otro gane o pierda.)

b) En cualquier juego la probabilidad de no empate (es decir, que gane A o B) es $q = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$ y la probabilidad de un empate es $p = 1 - q = \frac{1}{6}$. Entonces la probabilidad de 2 empates en 3 partidas es (vea el problema 1.37)

$$\binom{3}{2} p^2 q^{3-2} = 3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right) = \frac{5}{72}$$

c) $P(A \text{ y } B \text{ ganan alternadamente}) = P(\text{gana } A, \text{ después gana } B, \text{ después gana } A$ o B, después gana A, después gana $B)$

$$= P(A_1 \cap B_2 \cap A_3) + P(B_1 \cap A_2 \cap B_3)$$
$$= P(A_1)P(B_2)P(A_3) + P(B_1)P(A_2)P(B_3)$$
$$= \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) = \frac{5}{36}$$

d) $P(B \text{ gana por lo menos un juego}) = 1 - P(B \text{ no gana ningún juego})$

$$= 1 - P(B_1' \cap B_2' \cap B_3')$$
$$= 1 - P(B_1')\,P(B_2')\,P(B_3')$$
$$= 1 - \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right) = \frac{19}{27}$$

1.40. A y B juegan a lanzar alternadamente un par de dados. El primero que obtenga un total de 7 gana el juego. Encuentre la probabilidad de que *a*) el que lance primero los dados gane el juego, *b*) el segundo que lance los dados gane el juego.

a) La probabilidad de obtener 7 en un solo lanzamiento de un par de dados, suponiendo que los dados no estén cargados, es $1/6$ según lo hemos visto en el problema 1.9 y en la figura 1-9. Si se supone que A es el primero que lanza los dados, entonces en cualquiera de los siguientes casos mutuamente excluyentes, A ganará con las probabilidades asociadas indicadas:

1) A gana en el 1er. lanzamiento. Probabilidad $= \frac{1}{6}$.

2) A pierde en el 1er. lanzamiento, después pierde B, después gana A. Probabilidad $= \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)$.

3) A pierde en el 1er. lanzamiento, B pierde, A pierde, B pierde, A gana. Probabilidad $= \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)$.

.....

La probabilidad de que A gane es

$$\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \cdots$$
$$= \frac{1}{6}\left[1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \cdots\right] = \frac{1/6}{1 - (5/6)^2} = \frac{6}{11}$$

donde se ha utilizado el resultado 6 del apéndice A con $x = (5/6)^2$.

b) La probabilidad de que B gana el juego es similar

$$\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \cdots = \left(\frac{5}{6}\right)\left(\frac{1}{6}\right)\left[1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \cdots\right]$$
$$= \frac{5/36}{1 - (5/6)^2} = \frac{5}{11}$$

Por tanto, las probabilidades de que el primero en lanzar gane están 6 a 5. Observe que puesto que

$$\frac{6}{11} + \frac{5}{11} = 1$$

la probabilidad de empate es cero. Esto no sería verdad si el juego fuera limitado. Vea el problema 1.100.

1.41. Una máquina produce un total de 12 000 pernos al día, de los que en promedio 3% está defectuoso. Encontrar la probabilidad de que de 600 pernos elegidos al azar, 12 estén defectuosos.

De los 12 000 pernos, 3%, o 360, estén defectuosos y 11 640 no lo estén. Entonces:

$$\text{Probabilidad obtenida} = \frac{{}_{360}C_{12}\,{}_{11\,640}C_{588}}{{}_{12\,000}C_{600}}$$

1.42. Una caja contiene 5 canicas rojas y 4 blancas. De la caja se extraen sin reposición sucesivamente dos canicas, y se observa que la segunda es blanca. ¿Cuál es la probabilidad de que la primera también sea blanca?

Método 1
Si W_1 y W_2 son los eventos "blanca en la primera extracción", "blanca en la segunda extracción", respectivamente, lo que se busca es $P(W_1 \mid W_2)$. Esto está dado por

$$P(W_1 \mid W_2) = \frac{P(W_1 \cap W_2)}{P(W_2)} = \frac{(4/9)(3/8)}{4/9} = \frac{3}{8}$$

Método 2
Puesto que se sabe que la segunda es blanca, hay solamente 3 maneras para las 8 restantes en las que la primera pueda ser blanca, de modo que la probabilidad es $3/8$.

1.43. Las probabilidades que un esposo y una esposa sobrevivan 20 años son 0.8 y 0.9, respectivamente. Encuentre la probabilidad de que dentro de 20 años *a*) vivan ambos, *b*) ninguno de los dos viva, *c*) por lo menos uno esté vivo.

Sean H y W los eventos de que el esposo y la esposa, respectivamente, estén vivos dentro de 20 años. Entonces $P(H) = 0.8$, $P(W) = 0.9$. Se supone que los eventos H y W son independientes, lo que puede o no ser razonable.

a) $P(\text{ambos estén vivos}) = P(H \cap W) = P(H)P(W) = (0.8)(0.9) = 0.72$.

b) $P(\text{ninguno esté vivo}) = P(H' \cap W') = P(H')\,P(W') = (0.2)(0.1) = 0.02$.

c) $P(\text{por lo menos uno esté vivo}) = 1 - P\,(\text{ninguno esté vivo}) = 1 - 0.02 = 0.98$.

1.44. Una secretaria ineficiente pone al azar n cartas distintas en n sobres etiquetados con las diferentes direcciones. Encuentre la probabilidad de que por lo menos una de las cartas llegue al destino apropiado.

Sean $A_1, A_2, \ldots, A_n$ los eventos que la primera, segunda, $\ldots$, n-ésima carta esté en el sobre correcto. Entonces el evento de que por lo menos una carta esté en el sobre correcto es $A_1 \cup A_2 \cup \cdots \cup A_n$ y lo que se busca es $P(A_1 \cup A_2 \cup \cdots \cup A_n)$. Generalizando los resultados (10) y (11), página 6, se tiene

$$(1) \qquad P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum P(A_k) - \sum P(A_j \cap A_k) + \sum P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

donde $\sum P(A_k)$ es la suma de las probabilidades de las A_k de 1 a n, $\sum P(A_j \cap A_k)$ es la suma de las probabilidades $A_j \cap A_k$ con j y k de 1 a n y $k > j$, etc. Se tiene, por ejemplo, lo siguiente:

$$(2) \qquad P(A_1) = \frac{1}{n} \quad \text{y de manera similar} \quad P(A_k) = \frac{1}{n}$$

puesto que de los n sobres, sólo 1 tendrá la dirección apropiada. Además,

$$(3) \qquad P(A_1 \cap A_2) = P(A_1)\,P(A_2 \mid A_1) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)$$

puesto que, si la primera carta está en el sobre apropiado, entonces solamente 1 de los restantes $n - 1$ sobres será apropiado. De una manera similar se encuentra

$$(4) \qquad P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right)$$

etc., y por último

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\cdots\left(\frac{1}{1}\right) = \frac{1}{n!} \tag{5}$$

Ahora en $\sum P(A_j \cap A_k)$ hay $\binom{n}{2} = {}_nC_2$ términos todos con el mismo valor dado por (3). Similarmente en $\sum P(A_i \cap A_j \cap A_k)$ hay $\binom{n}{3} = {}_nC_3$ términos todos con el mismo valor dado por (4). Por tanto, la probabilidad buscada es

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) &= \binom{n}{1}\left(\frac{1}{n}\right) - \binom{n}{2}\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right) + \binom{n}{3}\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right) \\ &\quad - \cdots + (-1)^{n-1}\binom{n}{n}\left(\frac{1}{n!}\right) \\ &= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\frac{1}{n!} \end{aligned}$$

Del cálculo se sabe que (vea el apéndice A)

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

De manera que para $x = -1$

$$e^{-1} = 1 - \left(1 - \frac{1}{2!} + \frac{1}{3!} - \cdots\right)$$

o bien

$$1 - \frac{1}{2!} + \frac{1}{3!} - \cdots = 1 - e^{-1}$$

Se deduce que si n es grande, la probabilidad requerida es aproximadamente $1 - e^{-1} = 0.6321$. Esto significa que hay una buena posibilidad de que por lo menos 1 carta llegue al destino apropiado. Este resultado es interesante porque la probabilidad permanece casi constante para toda la $n > 10$. Por tanto, la probabilidad de que por lo menos 1 carta llegue a su destino es prácticamente la misma si n es 10 o 10 000.

1.45. Encontrar la probabilidad de que n personas seleccionadas al azar ($n \leq 365$) tengan n días de cumpleaños diferentes.

Supondremos que solamente hay 365 días en un año y que todos los cumpleaños son igualmente probables, supuestos que en la realidad no se satisfacen por completo.

La primera de las n personas tiene, por supuesto, algún día de cumpleaños con probabilidad $365/365 = 1$. Entonces, si la segunda debe tener un día de cumpleaños diferente, debe ocurrir en uno de los otros 364 días. Por tanto, la probabilidad de que la segunda persona tenga un día de cumpleaños diferente del de la primera es $364/365$. De manera semejante, la probabilidad de que la tercera persona tenga un día de cumpleaños diferente del de las dos primeras es $363/365$. Por último, la probabilidad de que la persona n-ésima tenga un día de cumpleaños diferente del de los demás es $(365 - n + 1)/365$. Por tanto, tenemos

$$\begin{aligned} P(\text{que los } n \text{ días de cumpleaños sean diferentes}) &= \frac{365}{365}\cdot\frac{364}{365}\cdot\frac{363}{365}\cdots\frac{365-n+1}{365} \\ &= \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)\cdots\left(1 - \frac{n-1}{365}\right) \end{aligned}$$

1.46. Determinar cuántas personas se requieren en el problema 1.45 para que la probabilidad de que los días de cumpleaños sean distintos sea menor que $1/2$.

Denotando la probabilidad dada por p y tomando logaritmos naturales, encontramos

$$\ln p = \ln\left(1 - \frac{1}{365}\right) + \ln\left(1 - \frac{2}{365}\right) + \cdots + \ln\left(1 - \frac{n-1}{365}\right) \tag{1}$$

pero del cálculo (apéndice A, fórmula 7) se sabe que

$$\ln(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots \tag{2}$$

de manera que (1) puede escribirse

$$\ln p = -\left[\frac{1 + 2 + \cdots + (n-1)}{365}\right] - \frac{1}{2}\left[\frac{1^2 + 2^2 + \cdots + (n-1)^2}{(365)^2}\right] - \cdots \tag{3}$$

Usando estos hechos para $n = 2, 3, \ldots$ (apéndice A, fórmulas 1 y 2)

$$1 + 2 + \cdots + (n-1) = \frac{n(n-1)}{2}, \qquad 1^2 + 2^2 + \cdots + (n-1)^2 = \frac{n(n-1)(2n-1)}{6} \tag{4}$$

para (3) se obtiene

$$\ln p = -\frac{n(n-1)}{730} - \frac{n(n-1)(2n-1)}{12(365)^2} - \cdots \tag{5}$$

Para n pequeña comparada con 365, por ejemplo, $n < 30$, el segundo término y los términos de orden superior en (5) son muy pequeños en comparación con el primer término, de manera que en este caso una buena aproximación es

$$\ln p = \frac{n(n-1)}{730} \tag{6}$$

Para $p = \frac{1}{2}$, $\ln p = -\ln 2 = -0.693$. Por tanto, se tiene

$$\frac{n(n-1)}{730} = 0.693 \qquad \text{o} \qquad n^2 - n - 506 = 0 \qquad \text{o} \qquad \quad - \quad + 22) = 0 \tag{7}$$

de manera que $n = 23$. La conclusión es, por tanto, que si n es mayor que 23, se puede decir con mayor seguridad que por lo menos 2 personas cumplan años el mismo día.

PROBLEMAS SUPLEMENTARIOS

CÁLCULO DE PROBABILIDADES

1.47. Determine o estime la probabilidad p de cada uno de los eventos siguientes:

a) Al sacar una sola carta de una baraja ordinaria bien mezclada se obtenga un rey, as, sota de tréboles o reina de diamantes.

b) Se obtenga la suma 8 en un solo lanzamiento de un par de dados.

c) Se obtenga un tornillo no defectuoso después de haber examinado 600 tornillos y 12 estén defectuosos.

d) Se obtenga un 7 o un 11 en un solo lanzamiento de un par de dados.

e) En tres lanzamientos de una moneda legal se obtenga por lo menos una cara.

1.48. Un experimento consiste en extraer tres cartas, una tras otra, de una baraja ordinaria bien mezclada. Sean A_1 el evento "rey en la primera extracción", A_2 el evento "rey en la segunda extracción" y A_3 el evento "rey en la tercera extracción". Exprese en palabras el significado de cada uno de los incisos siguientes:

a)

1.51. Repita el problema 1.50 sin reponer cada canica después de haber sido extraída.

PROBABILIDAD CONDICIONAL Y EVENTOS INDEPENDIENTES

1.52. Una caja contiene 2 canicas rojas y 3 azules. Encuentre la probabilidad de que si se extraen dos canicas al azar (sin reposición), *a*) las dos sean azules, *b*) las dos sean rojas, *c*) una sea roja y otra azul.

1.53. Encuentre la probabilidad de extraer al azar 3 ases de una baraja ordinaria de 52 cartas si las cartas *a*) se reponen, *b*) no se reponen.

1.54. Si en una familia hay 2 menores de edad, ¿cuál es la probabilidad de que uno sea varón, o de que los dos sean varones?

1.55. La caja *I* contiene 3 bolas rojas y 5 blancas, mientras que la caja *II* contiene 4 bolas rojas y 2 blancas. Se extrae una bola al azar de la primera caja y se coloca en la segunda caja sin observar su color. Después se extrae una bola de la segunda caja. Encuentre la probabilidad de que sea blanca.

TEOREMA O REGLA DE BAYES

1.56. Una caja contiene 3 canicas azules y 2 rojas mientras que otra caja contiene 2 canicas azules y 5 rojas. Una canica extraída al azar de una de las cajas resulta ser azul. ¿Cuál es la probabilidad que provenga de la primera caja?

1.57. Cada uno de tres joyeros idénticos tiene dos cajones. En cada cajón del primer joyero hay un reloj de oro. En cada cajón del segundo joyero hay un reloj de plata. En un cajón del tercer joyero hay un reloj de oro mientras que en el otro hay un reloj de plata. Si se selecciona un joyero al azar, se abre uno de los cajones y se encuentra que contiene un reloj de plata, ¿cuál es la probabilidad de que el otro cajón tenga un reloj de oro?

1.58. La urna *I* tiene 2 bolas blancas y 3 negras; la urna *II*, 4 blancas y 1 negra; y la urna *III*, 3 blancas y 4 negras. Se selecciona una urna al azar y una bola extraída al azar resulta ser blanca. Encuentre la probabilidad que se haya seleccionado la urna *I*.

ANÁLISIS COMBINATORIO, CONTEO Y DIAGRAMAS DE ÁRBOL

1.59. Una moneda se lanza 3 veces. Utilice un diagrama del árbol para determinar las diversas posibilidades que pueden presentarse.

1.60. Tres cartas se extraen al azar (sin reposición) de una baraja ordinaria de 52 cartas. Encuentre el número de maneras en las que se pueden extraer *a*) un diamante y un trébol y un corazón uno tras otro, *b*) dos corazones y después un trébol o una pica.

1.61. ¿De cuántas maneras pueden colocarse 3 monedas diferentes en dos monederos?

PERMUTACIONES

1.62. Evalúe *a*) ${}_4P_2$, *b*) ${}_7P_5$, *c*) ${}_{10}P_3$.

1.63. ¿Para qué valores de n es ${}_{n+1}P_3 = {}_nP_4$?

1.64. ¿De cuántas maneras puede sentarse a 5 personas en un sofá que sólo tiene 3 asientos?

1.65. ¿De cuántas maneras pueden ordenarse 7 libros en un anaquel si *a*) pueden ordenarse en cualquier orden, *b*) 3 libros determinados deben estar siempre juntos, *c*) dos libros determinados deben estar en los extremos.

1.66. ¿Cuántos números que consten de cinco dígitos diferentes cada uno pueden hacerse con los dígitos 1,2,3, . . . , 9 si *a*) los números deben ser impares, *b*) los primeros dos dígitos de cada número son pares?

1.67. Resuelva el problema 1.66 si se permite que los dígitos se repitan.

1.68. ¿Cuántos números diferentes de tres dígitos pueden hacerse con 3 cuatros, 4 dos y 2 tres?

1.69. ¿De cuántas maneras puede sentarse a 3 hombres y 3 mujeres a una mesa redonda si *a*) no se impone ninguna restricción, *b*) 2 mujeres determinadas no deben sentarse juntas, *c*) cada mujer debe estar entre 2 hombres?

COMBINACIONES

1.70. Evalúe *a*) ${}_5C_3$, *b*) ${}_8C_4$, *c*) ${}_{10}C_8$.

1.71. ¿Para qué valores de n es $3 \cdot {}_{n+1}C_3 = 7 \cdot {}_nC_2$?

1.72. ¿De cuántas maneras pueden elegirse 6 preguntas de un conjunto de 10 preguntas?

1.73. ¿Cuántos comités diferentes de 3 hombres y 4 mujeres pueden formarse a partir de 8 hombres y 6 mujeres?

1.74. ¿De cuántas maneras pueden ser seleccionados 2 hombres, 4 mujeres, 3 muchachos y 3 muchachas a partir de 6 hombres, 8 mujeres, 4 muchachos y 5 muchachas si *a*) no se impone ninguna restricción, *b*) un hombre y una mujer determinados deben ser seleccionados?

1.75. ¿De cuántas maneras se puede dividir un grupo de 10 personas en *a*) dos grupos que consten de 7 y 3 personas, *b*) tres grupos que consten de 5, 3 y 2 personas?

1.76. A partir de 5 especialistas en estadística y 6 economistas, debe formarse un comité que conste de 3 especialistas en estadística y 2 economistas. ¿Cuántos comités diferentes pueden formarse si *a*) no se impone ninguna restricción, *b*) dos determinados especialistas en estadística deben estar en el comité, *c*) un determinado economista no puede estar en el comité?

1.77. Encuentre el número *a*) de combinaciones y *b*) de permutaciones de 4 letras cada una que pueden hacerse con las letras de la palabra *Tennessee*.

COEFICIENTES BINOMIALES

1.78. Calcule *a*) ${}_6C_3$, *b*) $\binom{11}{4}$, *c*) $({}_8C_2)({}_4C_3)/{}_{12}C_5$.

1.79. Desarrolle *a*) $(x + y)^6$, *b*) $(x - y)^4$, *c*) $(x - x^{-1})^5$, *d*) $(x^2 + 2)^4$.

1.80. Encuentre el coeficiente de x en $\left(x + \frac{2}{x}\right)^9$.

PROBABILIDAD USANDO ANÁLISIS COMBINATORIO

1.81. Encuentre la probabilidad de anotar un total de 7 puntos *a*) una vez, *b*) por lo menos una vez, *c*) dos veces, en 2 lanzamientos de un par de dados no cargados.

1.82. De una baraja de 52 cartas bien mezcladas se extraen sucesivamente dos cartas. Encuentre la probabilidad de que *a*) la primera carta no sea un diez de tréboles o un as; *b*) la primera carta sea un as pero la segunda no; *c*) por lo menos una carta sea un diamante; *d*) las cartas no sean del mismo palo; *e*) no más de 1 carta sea una carta de figura (sota, reina, rey); *f*) la segunda carta no sea una carta de figura; *g*) la segunda carta no sea una carta de figura dado que la primera fue una carta de figura; *h*) las cartas sean cartas de figura o picas o ambas.

1.83. Una caja contiene 9 boletos numerados del 1 al 9 inclusive. Si se extraen 3 boletos de la caja, uno cada vez, encuentre la probabilidad de que sean alternadamente non, par, non, o par, non, par.

1.84. Las probabilidades a favor de que *A* gane un juego de ajedrez contra *B* son 3:2. Si se efectúan 3 juegos, ¿cuáles son las probabilidades *a*) a favor de que *A* gane por lo menos 2 de los 3 juegos, *b*) en contra de que *A* pierda los primeros 2 juegos contra *B*?

1.85. En el juego de *naipes*, a cada uno de los 4 jugadores se le dan 13 cartas de una baraja ordinaria bien mezclada de 52 cartas. Encuentre la probabilidad de que uno de los jugadores (por ejemplo, el más viejo) obtenga *a*) 7 diamantes, 2 tréboles, 3 corazones y 1 pica; *b*) un palo completo.

1.86. Una urna contiene 6 canicas rojas y 8 azules. Cinco canicas se extraen al azar sin reposición. Encuentre la probabilidad que 3 sean rojas y 2 azules.

1.87. *a*) Encuentre la probabilidad de obtener la suma 7 en por lo menos 1 de 3 lanzamientos de un par de dados no cargados. *b*) ¿Cuántos lanzamientos son necesarios para que la probabilidad en *a*) sea mayor a 0.95?

1.88. Tres cartas se extraen de una baraja ordinaria de 52 cartas. Encuentre la probabilidad de que *a*) todas las cartas sean de un palo, *b*) se extraigan por lo menos 2 ases.

1.89. Encuentre la probabilidad de que a un jugador de naipes que le dan 13 cartas, 9 sean de un mismo palo.

PROBLEMAS DIVERSOS

1.90. Un espacio muestral consta de 3 puntos muestrales con probabilidades dadas por $2p$, p^2 y $4p - 1$, respectivamente. Encuentre el valor de p.

1.91. ¿Cuántas palabras se pueden hacer a partir de 5 letras si *a*) todas las letras son diferentes, *b*) 2 letras son idénticas, *c*) todas las letras son diferentes pero 2 letras particulares no pueden estar adyacentes?

1.92. Cuatro números enteros se eligen al azar de entre el 0 y el 9, inclusive. Encuentre la probabilidad de que *a*) todos sean diferentes, *b*) no más de 2 sean iguales.

1.93. Un par de dados se lanza repetidamente. Encuentre la probabilidad que en el sexto lanzamiento se obtenga un 11 por primera vez.

1.94. ¿Cuál es el menor número de lanzamientos necesario en el problema 1.93 de modo que la probabilidad de obtener un 11 sea mayor a *a*) 0.5, *b*) 0.95?

1.95. Encuentre la probabilidad de que en un juego de póquer se obtenga *a*) un *royal flush*, que consta de diez, sota, reina, rey y as de un solo palo; *b*) *casa completa*, que consta de 3 cartas de un valor nominal y 2 de otro (como, por ejemplo, 3 dieces y 2 sotas); *c*) todas las cartas diferentes; *d*) 4 ases.

1.96. La probabilidad de que un hombre le dé al blanco es $\frac{2}{3}$. Si tira al blanco hasta que le da por primera vez, encuentre la probabilidad de que necesite 5 tiros para darle al blanco.

1.97. *a*) Un anaquel contiene 6 compartimientos separados. ¿De cuántas maneras se pueden colocar 4 canicas indistinguibles en los compartimientos? *b*) Repita el problema si hay n compartimientos y r canicas. Este tipo de problema se presenta en la física en relación con la *estadística de Bose-Einstein.*

1.98. *a*) Un estante contiene 6 compartimientos separados. ¿De cuántas maneras se pueden colocar 12 canicas indistinguibles en los compartimientos de modo que no haya ninguno vacío? *b*) Trabaje el problema si hay n compartimientos y r canicas donde $r > n$. Este tipo de problema se presenta en la física en relación con la *estadística de Fermi-Dirac*.

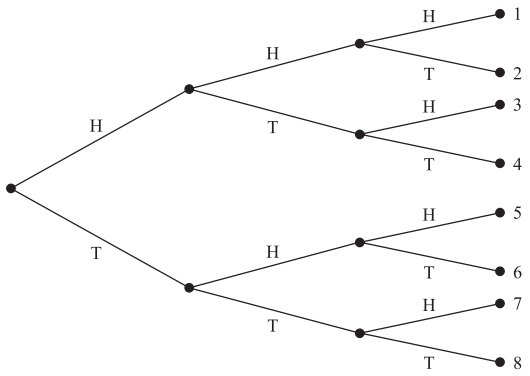
1.99. Un jugador de póquer tiene las cartas 2, 3, 4, 6, 8. Desea desechar el 8 y sustituirlo por otra carta que espera sea un 5 (en cuyo caso obtiene un "escalera cerrada"). ¿Cuál es la probabilidad de que tenga éxito si se supone que los otros tres jugadores juntos tienen *a*) 1 cinco, *b*) 2 cincos, *c*) 3 cincos, *d*) ningún cinco? ¿Puede resolverse el problema si el número de cincos en manos de los otros jugadores es desconocido? Explique.

1.100. Resolver el problema 1.40 si el juego está limitado a 3 lanzamientos.

1.101. Halle la probabilidad de que en un juego de naipes en que se reciben 13 cartas *a*) 2 jugadores, *b*) 3 jugadores, *c*) los cuatro jugadores tengan un palo completo.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

1.59.



1.60. *a*) $13 \times 13 \times 13$ *b*) $13 \times 12 \times 26$ **1.61.** 8 **1.62.** *a*) 12 *b*) 2 520 *c*) 720

1.63. $n = 5$ **1.64.** 60 **1.65.** *a*) 5 040 *b*) 720 *c*) 240 **1.66.** *a*) 8 400 *b*) 2 520

1.67. *a*) 32 805 *b*) 11 664 **1.68.** 26 **1.69.** *a*) 120 *b*) 72 *c*) 12

1.70. *a*) 10 *b*) 70 *c*) 45 **1.71.** $n = 6$ **1.72.** 210 **1.73.** 840

1.74. *a*) 42 000 *b*) 7 000 **1.75.** *a*) 120 *b*) 2 520 **1.76.** *a*) 150 *b*) 45 *c*) 100

1.77. *a*) 17 *b*) 163 **1.78.** *a*) 20 *b*) 330 *c*) 14/99

1.79. *a*) $x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^3 + 6xy^5 + y^6$

b) $x^4 - 4x^3y + 6x^2y^2 - 4xy^3 + y^4$

c) $x^5 - 5x^3 + 10x - 10x^{-1} + 5x^{-3} - x^{-5}$

d) $x^8 + 8x^6 + 24x^4 + 32x^2 + 16$

1.80. 2 016 **1.81.** *a*) 5/18 *b*) 11/36 *c*) 1/36

1.82. *a*) 47/52 *b*) 16/221 *c*) 15/34 *d*) 13/17 *e*) 210/221 *f*) 10/13 *g*) 40/51 *h*) 77/442

1.83. 5/18 **1.84.** *a*) 81 : 44 *b*) 21 : 4

1.85. *a*) $({}_{13}C_7)({}_{13}C_2)({}_{13}C_3)({}_{13}C_1)/{}_{52}C_{13}$ *b*) $4/{}_{52}C_{13}$ **1.86.** $({}_6C_3)({}_8C_2)/{}_{14}C_5$

1.87. *a*) 91/216 *b*) por lo menos 17 **1.88.** *a*) $4 \cdot {}_{13}C_{3//52}C_3$ *b*) $({}_4C_2 \cdot {}_{48}C_1 + {}_4C_3)/{}_{52}C_3$

1.89. $4({}_{13}C_9)({}_{39}C_4)/{}_{52}C_{13}$ **1.90.** $\sqrt{11} - 3$ **1.91.** *a*) 120 *b*) 60 *c*) 72

1.92. *a*) 63/125 *b*) 963/1 000 **1.93.** 1 419 857/34 012 224 **1.94.** *a*) 13 *b*) 53

1.95. *a*) $4/{}_{52}C_5$ *b*) $(13)(2)(4)(6)/{}_{52}C_5$ *c*) $4^5\,({}_{13}C_5)/{}_{52}C_5$ *d*) $(5)(4)(3)(2)/(52)(51)(50)(49)$

1.96. 2/243 **1.97.** *a*) 126 *b*) $_{n+r-1}C_{n-1}$ **1.98.** *a*) 462 *b*) $_{r-1}C_{n-1}$

1.99. *a*) 3/32 *b*) 1/16 *c*) 1/32 *d*) 1/8

1.100. prob. *A* gane = 61/216, prob. *B* gane = 5/36, prob. empate = 125/216

1.101. *a*) $12/(_{52}C_{13})(_{39}C_{13})$ *b*) $24/(_{52}C_{13})(_{39}C_{13})(_{26}C_{13})$

PROBLEMAS APORTADOS

1. En una escuela preparatoria que tiene 250 alumnos habrá votaciones para elegir a la Sociedad de alumnos, y van a formarse dos planillas. Cada planilla debe tener cinco integrantes. ¿De cuántas formas diferentes se pueden seleccionar estas planillas?

2. El departamento de control de calidad de una fábrica de envases de vidrio detectó tres tipos de defectos en el producto, los que clasificó como A, B y C. Se inspeccionó un lote de 300 envases, obteniéndose lo siguiente: 50 envases tenían defectos tipo A; 45 del tipo B; 40 del tipo C; 35 de los tipos A y B; 30 de los tipos A y C; 25 de los tipos B y C, y 250 no tenían ningún tipo de defecto. ¿Cuántos envases tenían los tres tipos de defectos?

3. En la siguiente tabla se muestra el consumo bimestral de energía eléctrica en KWH (kilowatts-hora) y el importe pagado para cada bimestre, por un periodo de dos años.

Bimestre	KWH	Importe (pesos)	Bimestre	KWH	Importe (pesos)
1/2011	299	484.00	1/2012	421	750.00
2/2011	314	489.00	2/2012	425	767.00
3/2011	321	516.00	3/2012	428	783.00
4/2011	410	805.00	4/2012	408	720.00
5/2011	404	685.00	5/2012	315	483.00
6/2011	401	679.00	6/2012	294	421.00

a) Expresar las cantidades bimestrales de consumo de energía eléctrica como porcentajes del consumo bimestral para cada año.
b) Graficar los porcentajes obtenidos en el inciso *a*).
c) Mediante Excel elaborar una gráfica de barras con columnas agrupadas.

4. Se tienen dos monedas y un dado no cargado, y se lanzan en el orden siguiente: moneda 1 (M1), dado (D), moneda 2 (M2); en el primer lanzamiento se desea obtener cara, número par y cruz. Determinar de cuántas maneras se puede obtener este resultado usando:

a) el principio fundamental de conteo
b) un diagrama de árbol

Propiedad	Unión	Intersección
Idempotencia	$A \cup A = A$	$A \cap A = A$
Conmutativa	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Asociación	$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$
Absorción	$A \cup (A \cap B) = A$	$A \cap (A \cup B) = A$
Distributiva	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Complementariedad	$A \cup A' = U$	$A \cap A' = \varnothing$

CAPÍTULO 2

VARIABLES ALEATORIAS

Suponga que a cada punto de un espacio muestral se le asigna un número. En este caso, se tiene una *función* definida sobre el espacio muestral. Esta función se denomina *variable aleatoria* (o *variable estocástica*), o, de manera más precisa, *función aleatoria* (*función estocástica*). Las variables aleatorias suelen denotarse con letras mayúsculas como, por ejemplo, X o Y. Por lo general, las variables aleatorias tienen un significado físico, geométrico o de algún otro tipo.

EJEMPLO 2.1 Suponga que una moneda se lanza dos veces de modo que el espacio muestral es $S = \{HH, HT, TH, TT\}$. Sea X el número de caras que se obtiene. A cada punto muestral se le puede asociar el número X como se indica en la tabla 2-1. Así, por ejemplo, en el caso de HH (es decir, 2 caras), $X = 2$, mientras que en el caso de TH (1 cara), $X = 1$. De lo anterior se deduce que X es una variable aleatoria.

Tabla 2-1

Punto muestral	HH	HT	TH	TT
X	2	1	1	0

Debe observarse que en este espacio muestral también pueden definirse muchas otras variables aleatorias, por ejemplo, el cuadrado del número de caras o el número de caras menos el número de cruces.

Una variable aleatoria que toma un número finito o un número infinito contable de valores (vea la página 4) se denomina *variable aleatoria discreta*, mientras que una que toma un número infinito no contable de valores se denomina *variable aleatoria no discreta*.

DISTRIBUCIONES DE PROBABILIDAD DISCRETAS

Sea X una variable aleatoria discreta, y suponga que sus valores posibles son $x_1, x_2, x_3, \ldots$, dispuestos en algún orden. También suponga que toma estos valores con las probabilidades dadas por

$$P(X = x_k) = f(x_k) \qquad k = 1, 2, \ldots \tag{1}$$

Es conveniente introducir la *función de probabilidad*, conocida también como *distribución de probabilidad*, dada por

$$P(X = x) = f(x) \tag{2}$$

Para $x = x_k$, esto se reduce a (1) mientras que para otros valores de x, $f(x) = 0$.

En general, $f(x)$ es una función de probabilidad si

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

donde la suma del inciso 2 se toma sobre todos los valores posibles de x.

EJEMPLO 2.2 Encuentre la función de probabilidad correspondiente a la variable aleatoria X del ejemplo 2.1. Si se supone que la moneda no está cargada, se tiene que

$$P(HH) = \frac{1}{4} \qquad P(HT) = \frac{1}{4} \qquad P(TH) = \frac{1}{4} \qquad P(TT) = \frac{1}{4}$$

En consecuencia,

$$P(X = 0) = P(TT) = \frac{1}{4}$$

$$P(X = 1) = P(HT \cup TH) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X = 2) = P(HH) = \frac{1}{4}$$

Por tanto, la función de probabilidad es la que se presenta en la tabla 2-2.

Tabla 2-2

x	0	1	2
$f(x)$	1/4	1/2	1/4

FUNCIONES DE DISTRIBUCIÓN DE VARIABLES ALEATORIAS

La *función de distribución acumulada*, o brevemente *función de distribución*, de una variable aleatoria X se define como

$$F(x) = P(X \leq x) \tag{3}$$

donde x es cualquier número real, es decir, $-\infty < x < \infty$.

La función de distribución $F(x)$ tiene las propiedades siguientes:

1. $F(x)$ es no decreciente [es decir, $F(x) \leq F(y)$ si $x \leq y$].
2. $\lim_{x \to -\infty} F(x) = 0$; $\lim_{x \to \infty} F(x) = 1$.
3. $F(x)$ es continua por la derecha [es decir, $\lim_{h \to 0^+} F(x + h) = F(x)$ para toda x].

FUNCIONES DE DISTRIBUCIÓN DE VARIABLES ALEATORIAS DISCRETAS

La función de distribución de una variable aleatoria discreta X puede obtenerse de su función de probabilidad observando que, para toda x en $(-\infty, \infty)$,

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u) \tag{4}$$

donde se suma sobre todos los valores u tomados por X para los que $u \leq x$.

Si X sólo toma un número finito de valores $x_1, x_2, \ldots, x_n$, entonces la función de distribución está dada por

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \leq x < \infty \end{cases} \tag{5}$$

EJEMPLO 2.3 *a*) Encuentre la función de distribución de la variable aleatoria X del ejemplo 2.2. *b*). Obtenga su gráfica.

a) La función de distribución es

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty \end{cases}$$

b) La gráfica de $F(x)$ se muestra en la figura 2-1.

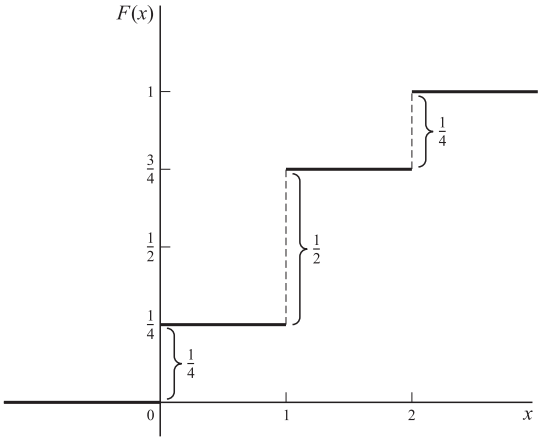


Figura 2-1

Se debe observar los siguientes aspectos que satisface, en general, la función de distribución anterior:

1. Las magnitudes de los saltos en 0, 1, 2 son $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ que son precisamente las probabilidades dadas en la tabla 2-2. Este hecho permite obtener la función de probabilidad a partir de la función de distribución.
2. Debido al aspecto de la gráfica de la figura 2-1, a menudo se le denomina *función escalonada*. El valor de la función es un número entero que se obtiene del escalón más alto; así, el valor en 1 es $\frac{3}{4}$ y no $\frac{1}{4}$. Esto se expresa en forma matemática diciendo que la función de distribución *es continua por la derecha* en 0, 1, 2.
3. Al desplazarse de izquierda a derecha (es decir, al *subir la escalera*), la función de distribución permanece igual o crece, es decir, toma valores desde 0 hasta 1. Debido a ello, se dice que es una *función monótona creciente*.

De acuerdo con las observaciones anteriores y de las propiedades de las funciones de distribución es claro que la función de probabilidad de una variable aleatoria discreta puede obtenerse a partir de la función de distribución si se toma en cuenta que

$$f(x) = F(x) - \lim_{u \to x^-} F(u). \tag{6}$$

VARIABLES ALEATORIAS CONTINUAS

Se dice que una variable aleatoria no discreta X es *absolutamente continua*, o simplemente *continua*, si su función de distribución puede representarse como

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du \qquad (-\infty < x < \infty) \tag{7}$$

donde la función $f(x)$ tiene las propiedades

1. $f(x) \geq 0$
2. $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$

De lo anterior se concluye que si X es una variable aleatoria continua, entonces la probabilidad que X asuma cualquier valor particular es cero, mientras que la *probabilidad de intervalo* de que X esté *entre dos valores diferentes*, por ejemplo, a y b, está dada por

$$P(a < X < b) = \int_{a}^{b} f(x)\,dx \tag{8}$$

EJEMPLO 2.4 Si de un grupo grande de varones adultos se selecciona al azar un individuo, la probabilidad de que su estatura X sea exactamente 68 pulgadas (es decir, 68.000… pulgadas) será cero. Sin embargo, existe una probabilidad mayor que cero de que X esté entre 67.000… pulgadas y 68.500… pulgadas, por ejemplo.

Una función $f(x)$ que satisface los requisitos anteriores se llama *función de probabilidad* o *distribución de probabilidad* de una variable aleatoria continua, aunque a menudo suele llamársele *función de densidad de probabilidad* o simplemente *función de densidad*. Toda función $f(x)$ que satisfaga las propiedades 1 y 2 dadas antes será de manera automática una función de densidad, y las probabilidades buscadas pueden obtenerse a partir de (8).

EJEMPLO 2.5 *a*) Encuentre una constante c tal que la función

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{si no es así} \end{cases}$$

sea una función de densidad y *b*) calcule $P(1 < x < 2)$.

a) Como $f(x)$ satisface la propiedad 1 si $c \geq 0$, debe satisfacer la propiedad 2 para ser una función de densidad. Ahora

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^3 cx^2\,dx = \frac{cx^3}{3}\bigg|_0^3 = 9c$$

y como esto debe ser igual a 1, se tiene $c = 1/9$.

b)
$$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \frac{x^3}{27}\bigg|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

En caso de que $f(x)$ sea continua, lo que se supondrá, a menos que se indique otra cosa, la probabilidad de que X sea igual a cualquier valor particular es cero. En tal caso puede sustituirse cualquiera de los signos, $<$ o $\leq$, o ambos en (8). Así, en el ejemplo 2.5,

$$P(1 \leq X \leq 2) = P(1 \leq X < 2) = P(1 < X \leq 2) = P(1 < X < 2) = \frac{7}{27}$$

EJEMPLO 2.6 *a*) Encuentre la función de distribución de la variable aleatoria del ejemplo 2.5. *b*) Utilice el resultado de *a*) para encontrar $P(1 < x \leq 2)$.

a) Se tiene

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

Si $x < 0$, entonces $F(x) = 0$. Si $0 \leq x < 3$, entonces

$$F(x) = \int_0^x f(u)\,du = \int_0^x \frac{1}{9}u^2\,du = \frac{x^3}{27}$$

Si $x \geq 3$, entonces

$$F(x) = \int_0^3 f(u)\,du + \int_3^x f(u)\,du = \int_0^3 \frac{1}{9}u^2\,du + \int_3^x 0\,du = 1$$

Por tanto, la función de distribución que se busca es

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3/27 & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Observe que $F(x)$ aumenta de manera monótona desde 0 hasta 1 como lo requiere una función de distribución. También debe observarse que en este caso $F(x)$ es continua.

b) Se tiene

$$\begin{aligned} P(1 < X \leq 2) &= P(X \leq 2) - P(X \leq 1) \\ &= F(2) - F(1) \\ &= \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27} \end{aligned}$$

como en el ejemplo 2.5.

La probabilidad de que X esté entre x y $x + \Delta x$ está dada por

$$P(x \leq X \leq x + \Delta x) = \int_x^{x+\Delta x} f(u)\,du \tag{9}$$

de manera que si Δx es pequeño, se tiene aproximadamente

$$P(x \leq X \leq x + \Delta x) = f(x)\Delta x \tag{10}$$

De (7) se ve que diferenciando ambos lados

$$\frac{dF(x)}{dx} = f(x) \tag{11}$$

en todos los puntos en los que $f(x)$ es continua; es decir, la derivada de la función de distribución es la función de densidad.

Debe hacerse notar que existen variables aleatorias que no son ni discretas ni continuas. Puede demostrarse que la variable aleatoria X con la de función de distribución siguiente es un ejemplo de este tipo de variables.

$$F(x) = \begin{cases} 0 & x < 1 \\ \dfrac{x}{2} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

Con objeto de obtener (11), se emplea la propiedad básica

$$\frac{d}{dx}\int_a^x f(u)\,du = f(x) \tag{12}$$

que es una versión del teorema fundamental del cálculo.

INTERPRETACIONES GRÁFICAS

Si $f(x)$ es la función de densidad de una variable aleatoria X, entonces $y = f(x)$ puede representarse en forma gráfica por medio de una curva como la que se muestra en la figura 2-2. Como $f(x) \geq 0$, la curva no puede encontrarse más abajo del eje x. Toda el área limitada por la curva y el eje x debe valer 1 por la propiedad 2 de la página 36. Desde un punto de vista geométrico, la probabilidad de que X esté entre a y b, es decir, $P(a < X < b)$, está representada por el área que se muestra sombreada, en la figura 2-2.

La función de distribución $F(x) = P(X \leq x)$ es una función monótona creciente que crece de 0 a 1 y que se representa por una curva como en la figura 2-3.

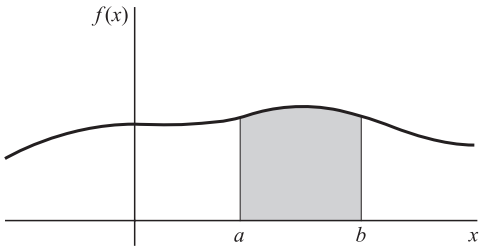


Figura 2-2

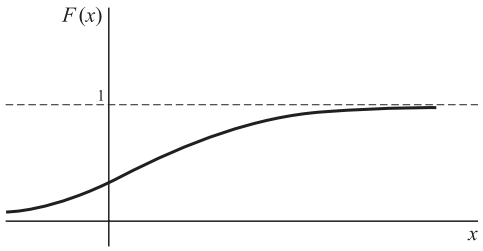


Figura 2-3

DISTRIBUCIONES CONJUNTAS

Es fácil generalizar las ideas anteriores a dos o más variables aleatorias. Se considerará el caso típico de dos variables aleatorias, ya sea que ambas sean discretas o ambas continuas. En los casos en que una variable es discreta y la otra es continua, es fácil hacer las modificaciones adecuadas. También puede hacerse la generalización a más de dos variables.

1. CASO DISCRETO. Si X y Y son dos variables aleatorias discretas, la *función de probabilidad conjunta* de X y Y se define como

$$P(X = x, Y = y) = f(x, y) \tag{13}$$

donde 1. $f(x, y) \geq 0$

2. $\sum_x \sum_y f(x, y) = 1$

es decir, la suma sobre todos los valores de x y y es 1.

Suponga que X puede tomar cualquiera de los m valores $x_1, x_2, \ldots, x_m$ y que Y puede tomar cualquiera de los n valores $y_1, y_2, \ldots, y_n$. En consecuencia, la probabilidad del evento $X = x_j$ y $Y = y_k$ está dada por

$$P(X = x_j, Y = y_k) = f(x_j, y_k) \tag{14}$$

Una función de probabilidad conjunta de X y Y puede representarse por medio de una *tabla de probabilidad conjunta* como la tabla 2-3. La probabilidad de que $X = x_j$ se obtiene mediante la suma de todas las entradas del renglón correspondiente a x_j y está dada por

$$P(X = x_j) = f_1(x_j) = \sum_{k=1}^{n} f(x_j, y_k) \tag{15}$$

Tabla 2-3

X \ Y	y_1	y_2	$\ldots$	y_n	Totales ↓
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\ldots$	$f(x_1, y_n)$	$f_1(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\ldots$	$f(x_2, y_n)$	$f_1(x_2)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	$\ldots$	$f(x_m, y_n)$	$f_1(x_m)$
Totales→	$f_2(y_1)$	$f_2(y_2)$	$\ldots$	$f_2(y_n)$	1 ← Gran total

Para $j = 1, 2, \ldots, m$, estas probabilidades corresponden a las entradas totales en la columna o margen del extremo derecho de la tabla 2-3. De manera similar, la probabilidad de que $Y = y_k$ se obtiene sumando todas las entradas de la columna correspondiente a y_k y está dada por

$$P(Y = y_k) = f_2(y_k) = \sum_{j=1}^{m} f(x_j, y_k) \tag{16}$$

Para $k = 1, 2, \ldots, n$ estas probabilidades corresponden a las entradas totales en el renglón o margen inferior de la tabla 2-3.

Dado que las probabilidades (15) y (16) se obtienen de los márgenes de la tabla, a $f_1(x_j)$ y a $f_2(y_k)$ [o simplemente $f_1(x)$ y $f_2(y)$] suele conocérseles como *funciones de probabilidad marginal* de X y Y, respectivamente.

También hay que observar que

$$\sum_{j=1}^{m} f_1(x_j) = 1 \quad \sum_{k=1}^{n} f_2(y_k) = 1 \tag{17}$$

lo cual puede escribirse como

$$\sum_{j=1}^{m}\sum_{k=1}^{n} f(x_j, y_k) = 1 \tag{18}$$

Esto simplemente dice que la probabilidad total de todas las entradas es 1. El *gran total* de 1 se muestra en la esquina inferior derecha de la tabla.

La *función de distribución conjunta* de X y Y se define como

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x}\sum_{v \leq y} f(u, v) \tag{19}$$

En la tabla 2-3, $F(x, y)$ es la suma de todas las entradas tales que $x_j \leq x$ y $y_k \leq y$.

2. CASO CONTINUO. Es fácil resolver, por analogía con el caso discreto, el caso en el que las dos variables son continuas mediante la sustitución de sumas por integrales. De esta manera, la *función de probabilidad conjunta* de las variables aleatorias X y Y (o, como se le denomina más comúnmente, la *función de densidad conjunta* de X y Y) se define como

1. $f(x, y) \geq 0$
2. $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$

En forma gráfica, $z = f(x, y)$ representa una superficie, como se indica en la figura 2-4, a la que se le denomina *superficie de probabilidad*. El volumen total que queda limitado por esta superficie y el plano xy es igual a 1, de acuerdo con la propiedad 2 antes citada. La probabilidad de que X esté entre a y b y Y esté entre c y d se muestra en forma gráfica por el volumen sombreado de la figura 2-4 y en forma matemática por

$$P(a < X < b, c < Y < d) = \int_{x=a}^{b}\int_{y=c}^{d} f(x, y)\,dx\,dy \tag{20}$$

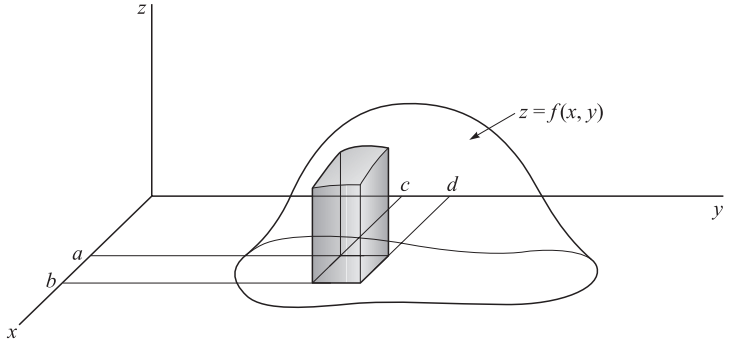


Figura 2-4

De manera más general, si A representa un evento, habrá una región $\mathcal{R}_A$ del plano xy que corresponda a ese evento. En este caso la probabilidad de A se encuentra integrando sobre $\mathcal{R}_A$, es decir,

$$P(A) = \iint_{\mathcal{R}_A} f(x, y)\,dx\,dy \tag{21}$$

En consecuencia, la *función de distribución conjunta* de X y Y se define como

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{y} f(u, v)\,du\,dv \tag{22}$$

Se deduce por analogía con (11), página 38, que

$$\frac{\partial^2 F}{\partial x\,\partial y} = f(x, y) \tag{23}$$

es decir, la función de densidad se obtiene mediante diferenciación de la función de distribución con respecto tanto a x como a y.

De acuerdo con (22) se obtiene

$$P(X \leq x) = F_1(x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv \tag{24}$$

$$P(Y \leq y) = F_2(y) = \int_{u=-\infty}^{\infty}\int_{v=-\infty}^{y} f(u, v)\,du\,dv \tag{25}$$

A (24) y (25) se les denomina *funciones de distribución marginal*, o simplemente *funciones de distribución*, de X y Y, respectivamente. A las derivadas de (24) y (25) respecto de x y y se les denomina *funciones de densidad marginal* o simplemente *funciones de densidad* de X y Y; además, están dadas por

$$f_1(x) = \int_{v=-\infty}^{\infty} f(x, v)\,dv \qquad f_2(y) = \int_{u=-\infty}^{\infty} f(u, y)\,du \tag{26}$$

VARIABLES ALEATORIAS INDEPENDIENTES

Suponga que X y Y son variables aleatorias discretas. Si los eventos $X = x$ y $Y = y$ son independientes para toda x y y, se dice que X y Y son *variables aleatorias independientes*. En este caso,

$$P(X = x, Y = y) = P(X = x)P(Y = y) \tag{27}$$

o, de manera equivalente,

$$f(x, y) = f_1(x)f_2(y) \tag{28}$$

A la inversa, si para toda x y y la función de probabilidad conjunta $f(x, y)$ puede expresarse como el producto de una función sólo de x por una función sólo de y (las cuales son las funciones de probabilidad marginal de X y Y), entonces X y Y son independientes. Pero, si $f(x, y)$ no puede expresarse de esta manera, entonces X y Y son *dependientes*.

Si X y Y son variables aleatorias continuas, se dice que son *variables aleatorias independientes* si los eventos $X \leq x$ y $Y \leq y$ son eventos independientes para toda x y y. En ese caso puede escribirse

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y) \tag{29}$$

o, lo que es equivalente,

$$F(x, y) = F_1(x)F_2(y) \tag{30}$$

donde $F_1(x)$ y $F_2(y)$ son las funciones de distribución (marginal) de X y Y, respectivamente. De manera inversa, X y Y son variables aleatorias independientes si para toda x y y, la función de distribución conjunta $F(x, y)$ puede expresarse como el producto de una función sólo de x y una función sólo de y (las cuales son las distribuciones marginales de X y Y, respectivamente). Pero si $F(x, y)$ no puede expresarse de esta manera, entonces X y Y son dependientes.

En el caso de variables aleatorias continuas independientes, la función de densidad conjunta $f(x, y)$ es también el producto de una función sólo de x, $f_1(x)$, por una función sólo de y, $f_2(y)$, y éstas son las funciones de densidad (marginal) de X y Y, respectivamente.

CAMBIO DE VARIABLES

Dadas las distribuciones de probabilidad de una o más variables aleatorias, es necesario hallar distribuciones de otras variables aleatorias que de alguna manera específica dependan de ellas. En los teoremas siguientes se presentan procedimientos para obtener estas distribuciones para los casos de variables discretas y continuas.

1. VARIABLES DISCRETAS

Teorema 2-1 Sea X una variable aleatoria discreta con función de probabilidad $f(x)$. Suponga que una variable aleatoria discreta U está definida en términos de X mediante $U = \phi(X)$, donde a cada valor de X le corresponde uno y sólo un valor de U, e inversamente, de manera que $X = \psi(U)$. En consecuencia, la función de probabilidad de U está dada por

$$g(u) = f[\psi(u)] \tag{31}$$

Teorema 2-2 Sean X y Y variables aleatorias discretas con la función de probabilidad conjunta $f(x, y)$. Suponga que dos variables aleatorias discretas U y V están definidas en términos de X y Y mediante $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$, donde a cada par de valores de X y Y corresponde uno y sólo un par de valores de U y V, y de manera inversa, de forma que $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$. Entonces, la función de probabilidad conjunta de U y V está dada por

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)] \tag{32}$$

2. VARIABLES CONTINUAS

Teorema 2-3 Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. Se define $U = \phi(X)$ donde $X = \psi(U)$ como en el teorema 2-1. En consecuencia, la densidad de probabilidad de U está dada por $g(u)$ donde

$$g(u)|du| = f(x)|dx| \tag{33}$$

o bien

$$g(u) = f(x)\left|\frac{dx}{du}\right| = f[\psi(u)]\left|\psi'(u)\right| \tag{34}$$

Teorema 2-4 Sean X y Y variables aleatorias continuas que tengan la función de probabilidad conjunta $f(x, y)$. Se define $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$ donde $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$ como en el teorema 2-2. Por tanto, la función de densidad conjunta de U y V está dada por $g(u,v)$ donde

$$g(u, v)|du\, dv| = f(x, y)|dx\, dy| \tag{35}$$

o bien

$$g(u, v) = f(x, y)\left|\frac{\partial(x, y)}{\partial(u, v)}\right| = f[\psi_1(u, v), \psi_2(u, v)]\left|J\right| \tag{36}$$

En (36) el *determinante jacobiano* o simplemente el *jacobiano* está dado por

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \tag{37}$$

DISTRIBUCIONES DE PROBABILIDAD DE FUNCIONES DE VARIABLES ALEATORIAS

En los teoremas 2-2 y 2-4 intervienen funciones de probabilidad conjunta de dos variables aleatorias. En la práctica, suele ser necesario hallar la distribución de probabilidad de alguna función determinada de varias variables aleatorias. En tales casos cualquiera de los teoremas siguientes suele ser útil.

Teorema 2-5 Sean X y Y variables aleatorias continuas y sea $U = \phi_1(X, Y)$, $V = X$ (la segunda elección es arbitraria). En consecuencia, la función de densidad de U es la densidad marginal obtenida de la densidad conjunta de U y V como en el teorema 2-4. En el caso de funciones de probabilidad de variables discretas es válido un resultado similar.

Teorema 2-6 Sea $f(x, y)$ la función de densidad conjunta de X y Y. Entonces, la función de densidad $g(u)$ de la variable aleatoria $U = \phi_1(X, Y)$ se encuentra diferenciando respecto a u la función de distribución

dada por

$$G(u) = P[\phi_1(X, Y) \leq u] = \iint_{\mathcal{R}} f(x, y)\, dx\, dy \tag{38}$$

donde $\mathcal{R}$ es la región en la que $\phi_1(x, y) \leq u$.

CONVOLUCIONES

Como consecuencia particular de los teoremas anteriores, se demuestra (vea el problema 2.23) que la función de densidad de la suma de dos variables aleatorias continuas X y Y, es decir $U = X + Y$, cuya función de densidad conjunta sea $f(x, y)$ está dada por

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx \tag{39}$$

En el caso especial en el que X y Y sean independientes, $f(x, y) = f_1(x)f_2(y)$, y (39) se reduce a

$$g(u) = \int_{-\infty}^{\infty} f_1(x)\, f_2(u - x)\, dx \tag{40}$$

a lo que se le conoce como *convolución* de f_1 y f_2, que se abrevia $f_1 * f_2$.

Las siguientes son algunas de las propiedades importantes de la convolución:

1. $f_1 * f_2 = f_2 * f_1$
2. $f_1 *(f_2 * f_3) = (f_1 * f_2) * f_3$
3. $f_1 *(f_2 + f_3) = f_1 * f_2 + f_1 * f_3$

Estos resultados muestran que f_1, f_2 y f_3 satisfacen las propiedades *conmutativa*, *asociativa* y *distributiva* del álgebra respecto de la operación de convolución.

DISTRIBUCIONES CONDICIONALES

Como ya se sabe, si $P(A) > 0$,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{41}$$

Si X y Y son variables aleatorias discretas y se tienen los eventos $(A: X = x)$, $(B: Y = y)$, entonces (41) se convierte en

$$P(Y = y \mid X = x) = \frac{f(x, y)}{f_1(x)} \tag{42}$$

donde $f(x, y) = P(X = x, Y = y)$ es la función de probabilidad conjunta y $f_1(x)$ es la función de probabilidad marginal de X. Se define

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{43}$$

y se le denomina *función de probabilidad condicional de Y dado X*. De manera similar, la función de probabilidad condicional de X dado Y es

$$f(x \mid y) \equiv \frac{f(x, y)}{f_2(y)} \tag{44}$$

Algunas veces $f(x \mid y)$ y $f(y \mid x)$ se denotarán $f_1(x \mid y)$ y $f_2(y \mid x)$, respectivamente.

Es fácil extender estas ideas al caso en el que X y Y sean variables aleatorias continuas. Por ejemplo, la *función de densidad condicional de Y dado X* es

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{45}$$

donde $f(x, y)$ es la función de densidad conjunta de X y Y, y $f_1(x)$ es la función de densidad marginal de X. Por ejemplo, usando (45) se puede determinar que la probabilidad de que Y se encuentra entre c y d dado que $x < X < x + dx$ es

$$P(c < Y < d \,|\, x < X < x + dx) = \int_c^d f(y \,|\, x)\, dy \tag{46}$$

También existe la generalización de estos resultados.

APLICACIONES A LA PROBABILIDAD GEOMÉTRICA

En probabilidad existen varios problemas que surgen de consideraciones geométricas o que tienen interpretaciones geométricas. Por ejemplo, suponga que se tiene un blanco que tiene la forma de una región plana de área K, con una porción de área K_1, como se muestra en la figura 2-5. En consecuencia, es razonable suponer que la probabilidad de acertar a la región de área K_1 es proporcional a K. De esta manera se define

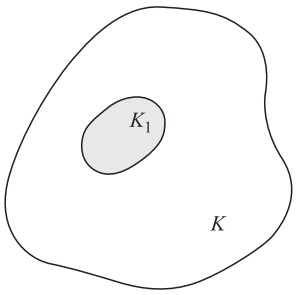


Figura 2-5

$$P(\text{acertar a la región de área } K_1) = \frac{K_1}{K} \tag{47}$$

donde se supone que la probabilidad de acertar al blanco es 1. Por supuesto, pueden hacerse otras suposiciones. Por ejemplo, habrá menos probabilidad de acertar a áreas exteriores. El tipo de suposición empleada define la función de distribución de probabilidad.

PROBLEMAS RESUELTOS

VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD DISCRETAS

2.1. Se lanza un par de dados no cargados, y se define X como la variable aleatoria que representa la suma de los puntos. Determinar la distribución de probabilidad de X.

En la figura 1-9, página 14, se dan los puntos muestrales del lanzamiento de un par de dados. La variable aleatoria X es la suma de las coordenadas de cada punto. Así, para (3, 2) se tiene que $X = 5$. Con base en el hecho de que los 36 puntos muestrales son igualmente probables, sabemos que cada punto muestral tiene probabilidad $1/36$, con esta información construimos la tabla 2-4. Por ejemplo, para $X = 5$ tenemos los puntos muestrales (1, 4), (2, 3), (3, 2), (4, 1), de manera que su probabilidad correspondiente es $4/36$.

Tabla 2-4

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

2.2. Determinar la distribución de probabilidad correspondiente a niños y niñas en familias de 3 hijos. Suponemos probabilidades iguales para niños y para niñas.

En el problema 1.37 se trató el caso de n ensayos mutuamente independientes, donde cada ensayo tenía sólo dos resultados posibles, A y A', con probabilidades respectivas p y $q = 1 - p$. Encontramos que la probabilidad de obtener exactamente x número de A en n ensayos es ${}_nC_x p^x q^{n-x}$. Este resultado lo aplicamos al problema presente, suponemos que los nacimientos sucesivos (los "ensayos") son independientes en lo que se refiere al sexo de los hijos. Por tanto, si A es el evento "un niño", $n = 3$ y p y $q = \frac{1}{2}$, tenemos

$$P(\text{exactamente un niño}) = P(X = x) = {}_3C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x} = {}_3C_x\left(\frac{1}{2}\right)^3$$

donde la variable aleatoria X representa el número de niños en la familia. (Observe que X está definida en el espacio muestral de 3 ensayos.) La función de probabilidad de X,

$$f(x) = {}_3C_x\left(\frac{1}{2}\right)^3$$

se presenta en la tabla 2-5.

Tabla 2-5

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

FUNCIONES DE DISTRIBUCIÓN DISCRETA

2.3. *a*) Encontrar la función de distribución $F(x)$ de la variable aleatoria X del problema 2.1 y *b*) graficar esta función de distribución.

a) Tenemos $F(x) = P(X \le x) = \sum_{u \le x} f(u)$. Entonces, de acuerdo con los resultados del problema 2.1 encontramos que

$$F(x) = \begin{cases} 0 & -\infty < x < 2 \\ 1/36 & 2 \le x < 3 \\ 3/36 & 3 \le x < 4 \\ 6/36 & 4 \le x < 5 \\ \vdots & \vdots \\ 35/36 & 11 \le x < 12 \\ 1 & 12 \le x < \infty \end{cases}$$

b) Vea la figura 2-6.

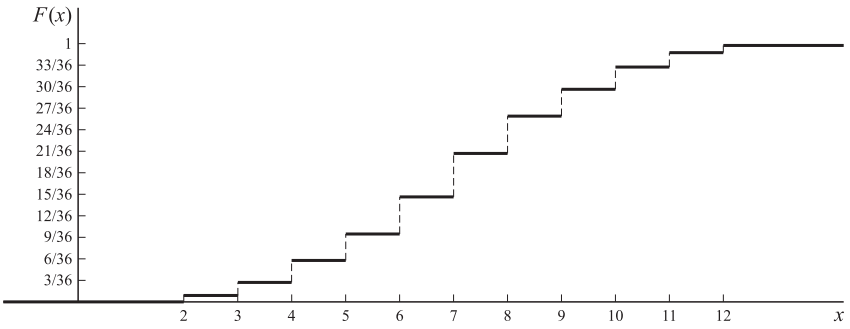


Figura 2-6

2.4. *a*) Encontrar la función de distribución $F(x)$ de la variable aleatoria X del problema 2.2 y *b*) graficar esta función de distribución.

a) Usando la tabla 2-5 del problema 2.2, obtenemos

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/8 & 0 \le x < 1 \\ 1/2 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & 3 \le x < \infty \end{cases}$$

b) En la figura 2-7 se muestra la gráfica de la función de distribución de *a*).

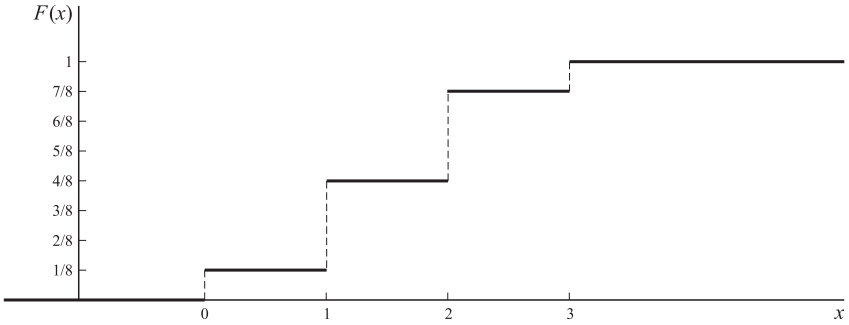


Figura 2-7

VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD CONTINUAS

2.5. Una variable aleatoria X tiene la función de densidad $f(x) = c/(x^2 + 1)$, donde $-\infty < x < \infty$. *a*) Encontrar el valor de la constante c. *b*) Encontrar la probabilidad de que X^2 esté entre $1/3$ y 1.

a) Se necesita tener $\int_{-\infty}^{\infty} f(x)\,dx = 1$, es decir,

$$\int_{-\infty}^{\infty} \frac{c\,dx}{x^2 + 1} = c \tan^{-1} x \Big|_{-\infty}^{\infty} = c\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1$$

de manera que $c = 1/\pi$.

b) Si $\frac{1}{3} \le X^2 \le 1$, entonces $\frac{\sqrt{3}}{3} \le X \le 1$ o bien $-1 \le X \le -\frac{\sqrt{3}}{3}$. Por tanto, la probabilidad que buscamos es

$$\frac{1}{\pi}\int_{-1}^{-\sqrt{3}/3} \frac{dx}{x^2 + 1} + \frac{1}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1} = \frac{2}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1}$$

$$= \frac{2}{\pi}\left[\tan^{-1}(1) - \tan^{-1}\left(\frac{\sqrt{3}}{3}\right)\right]$$

$$= \frac{2}{\pi}\left(\frac{\pi}{4} - \frac{\pi}{6}\right) = \frac{1}{6}$$

2.6. Encontrar la función de distribución correspondiente a la función de densidad del problema 2.5.

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \frac{1}{\pi}\int_{-\infty}^{x} \frac{du}{u^2 + 1} = \frac{1}{\pi}\left[\tan^{-1} u\Big|_{-\infty}^{x}\right]$$

$$= \frac{1}{\pi}[\tan^{-1} x - \tan^{-1}(-\infty)] = \frac{1}{\pi}\left[\tan^{-1} x + \frac{\pi}{2}\right]$$

$$= \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x$$

2.7. La función de distribución de una variable aleatoria X es

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Encontrar *a*) la función de densidad, *b*) la probabilidad de que $X > 2$ y *c*) la probabilidad de que $-3 < X \leq 4$.

a) $$f(x) = \frac{d}{dx}F(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x < 0 \end{cases}$$

b) $$P(X > 2) = \int_2^\infty 2e^{-2u}du = -e^{-2u}\Big|_2^\infty = e^{-4}$$

Otro método

Por definición, $P(X \leq 2) = F(2) = 1 - e^{-4}$. Por tanto,

$$P(X > 2) = 1 - (1 - e^{-4}) = e^{-4}$$

c) $$P(-3 < X \leq 4) = \int_{-3}^4 f(u)\,du = \int_{-3}^0 0\,du + \int_0^4 2e^{-2u}\,du$$
$$= -e^{-2u}\Big|_0^4 = 1 - e^{-8}$$

Otro método

$$P(-3 < X \leq 4) = P(X \leq 4) - P(X \leq -3)$$
$$= F(4) - F(-3)$$
$$= (1 - e^{-8}) - (0) = 1 - e^{-8}$$

DISTRIBUCIONES CONJUNTAS Y VARIABLES INDEPENDIENTES

2.8. La función de probabilidad conjunta de dos variables aleatorias discretas X y Y está dada por $f(x, y) = c(2x + y)$, donde x y y pueden tomar todos los valores enteros tales que $0 \leq x \leq 2$, $0 \leq y \leq 3$ y $f(x, y) = 0$ en los demás casos.

a) Encontrar el valor de la constante c.
b) Encontrar $P(X = 2, Y = 1)$.
c) Encontrar $P(X \geq 1, Y \leq 2)$.

a) En la figura 2-8 se muestran los puntos muestrales (x, y) para los que la probabilidades son distintas de cero. Las probabilidades que corresponden a estos puntos, están dadas por $c(2x + y)$, y se muestran en la tabla 2-6. Como el gran total, $42c$, debe ser igual a 1, se tiene $c = 1/42$.

Tabla 2-6

X \ Y	0	1	2	3	Totales ↓
0	0	c	$2c$	$3c$	$6c$
1	$2c$	$3c$	$4c$	$5c$	$14c$
2	$4c$	$5c$	$6c$	$7c$	$22c$
Totales →	$6c$	$9c$	$12c$	$15c$	$42c$

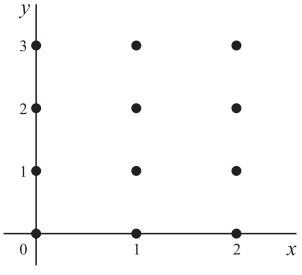


Figura 2-8

b) De acuerdo con la tabla 2-6 vemos que

$$P(X = 2, Y = 1) = 5c = \frac{5}{42}$$

c) En la tabla 2-6 vemos que

$$P(X \geq 1, Y \leq 2) = \sum_{x \geq 1} \sum_{y \leq 2} f(x, y)$$
$$= (2c + 3c + 4c) + (4c + 5c + 6c)$$
$$= 24c = \frac{24}{42} = \frac{4}{7}$$

como se indica mediante las entradas que se muestran sombreadas en la tabla.

2.9. Encontrar las funciones de probabilidad marginal *a*) de X y *b*) de Y para las variables aleatorias del problema 2.8.

a) La función de probabilidad marginal de X está dada por $P(X = x) = f_1(x)$ y lo obtenemos de los totales marginales de la columna del extremo derecho de la tabla 2-6. Ahí vemos que

$$P(X = x) = f_1(x) = \begin{cases} 6c = 1/7 & x = 0 \\ 14c = 1/3 & x = 1 \\ 22c = 11/21 & x = 2 \end{cases}$$

Verificación: $\frac{1}{7} + \frac{1}{3} + \frac{11}{21} = 1$

b) La función de probabilidad marginal de Y está dada por $P(Y = y) = f_2(y)$ y lo obtenemos de los totales marginales del último renglón inferior de la tabla 2-6, de donde vemos que

$$P(Y = y) = f_2(y) = \begin{cases} 6c = 1/7 & y = 0 \\ 9c = 3/14 & y = 1 \\ 12c = 2/7 & y = 2 \\ 15c = 5/14 & y = 3 \end{cases}$$

Verificación: $\frac{1}{7} + \frac{3}{14} + \frac{2}{7} + \frac{5}{14} = 1$

2.10. Mostrar que las variables aleatorias X y Y del problema 2.8 son dependientes.

Si las variables aleatorias X y Y fueran independientes, entonces, para toda x y y debería tenerse

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

Pero, como se ve de acuerdo con los problemas 2.8*b*) y 2.9,

$$P(X = 2, Y = 1) = \frac{5}{42} \qquad P(X = 2) = \frac{11}{21} \qquad P(Y = 1) = \frac{3}{14}$$

de manera que

$$P(X = 2, Y = 1) \neq P(X = 2)P(Y = 1)$$

Este resultado también es consecuencia del hecho de que la función de probabilidad conjunta $(2x + y)/42$ no puede expresarse como el producto de una función sólo de x por una función sólo de y.

2.11. La función de densidad conjunta de dos variables aleatorias continuas X y Y es

$$f(x, y) = \begin{cases} cxy & 0 < x < 4, 1 < y < 5 \\ 0 & \text{si no es así} \end{cases}$$

a) Encontrar el valor de la constante c.
b) Encontrar $P(1 < X < 2, 2 < Y < 3)$.
c) Encontrar $P(X \geq 3, Y \leq 2)$.

a) La probabilidad total debe ser igual a 1, es decir,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$$

Usando la definición de $f(x, y)$, la integral tiene el valor

$$\int_{x=0}^{4}\int_{y=1}^{5} cxy\,dx\,dy = c\int_{x=0}^{4}\left[\int_{y=1}^{5} xy\,dy\right]dx$$

$$= c\int_{z=0}^{4} \frac{xy^2}{2}\bigg|_{y=1}^{5} dx = c\int_{x=0}^{4}\left(\frac{25x}{2} - \frac{x}{2}\right)dx$$

$$= c\int_{x=0}^{4} 12x\,dx = c(6x^2)\Big|_{x=0}^{4} = 96c$$

Entonces, $96c = 1$ y $c = 1/96$.

b) Empleando el valor de c que se determinó en el inciso *a*), tenemos

$$P(1 < X < 2, 2 < Y < 3) = \int_{x=1}^{2}\int_{y=2}^{3} \frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=1}^{2}\left[\int_{y=2}^{3} xy\,dy\right]dx = \frac{1}{96}\int_{x=1}^{2} \frac{xy^2}{2}\bigg|_{y=2}^{3} dx$$

$$= \frac{1}{96}\int_{x=1}^{2} \frac{5x}{2}dx = \frac{5}{192}\left(\frac{x^2}{2}\right)\Big|_{1}^{2} = \frac{5}{128}$$

c)
$$P(X \geq 3, Y \leq 2) = \int_{x=3}^{4}\int_{y=1}^{2} \frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=3}^{4}\left[\int_{y=1}^{2} xy\,dy\right]dx = \frac{1}{96}\int_{x=3}^{4} \frac{xy^2}{2}\bigg|_{y=1}^{2} dx$$

$$= \frac{1}{96}\int_{x=3}^{4} \frac{3x}{2}dx = \frac{7}{128}$$

2.12. Encontrar la función de distribución marginal *a*) de X y *b*) de Y en el problema 2.11.

a) La función de distribución marginal de X si $0 \leq x < 4$ es

$$F_1(x) = P(X \leq x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv$$

$$= \int_{u=0}^{x}\int_{v=1}^{5} \frac{uv}{96}\,du\,dv$$

$$= \frac{1}{96}\int_{u=0}^{x}\left[\int_{v=1}^{5} uv\,dv\right]du = \frac{x^2}{16}$$

Para $x \geq 4$, $F_1(x) = 1$; para $x < 0$, $F_1(x) = 0$. Por tanto,

$$F_1(x) = \begin{cases} 0 & x < 0 \\ x^{2/16} & 0 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

Como $F_1(x)$ es continua en $x = 0$ y en $x = 4$, en la expresión anterior $<$ puede sustituirse por $\leq$.

b) La función de distribución marginal de Y si $1 \le y < 5$ es

$$F_2(y) = P(Y \le y) = \int_{u=-\infty}^{\infty} \int_{v=1}^{y} f(u, v)\, du\, dv$$

$$= \int_{u=0}^{4} \int_{v=1}^{y} \frac{uv}{96}\, du\, dv = \frac{y^2 - 1}{24}$$

Para $y \ge 5$, $F_2(y) = 1$; para $y < 1$, $F_2(y) = 0$. Por tanto,

$$F_2(y) = \begin{cases} 0 & y < 1 \\ (y^2 - 1)/24 & 1 \le y < 5 \\ 1 & y \ge 5 \end{cases}$$

Como $F_2(y)$ es continua en $y = 1$ y en $y = 5$, en la expresión anterior $<$ puede sustituirse por $\le$.

2.13. Encontrar la función de distribución conjunta de las variables X y Y del problema 2.11.

De acuerdo con el problema 2.11 se ve que la función de densidad conjunta de X y Y puede escribirse como el producto de una función sólo de x y una función sólo de y. En efecto: $f(x, y) = f_1(x)f_2(y)$, donde

$$f_1(x) = \begin{cases} c_1 x & 0 < x < 4 \\ 0 & \text{si no es así} \end{cases} \qquad f_2(y) = \begin{cases} c_2 y & 1 < y < 5 \\ 0 & \text{si no es así} \end{cases}$$

y $c_1 c_2 = c = 1/96$. Se infiere que X y Y son independientes, de manera que su función de distribución conjunta está dada por $F(x, y) = F_1(x)F_2(y)$. Las distribuciones marginales $F_1(x)$ y $F_2(y)$ se determinaron en el problema 2.12 y en la figura 2-9 se muestra la definición por partes resultante de $F(x, y)$.

2.14. En el problema 2.11 encontrar $P(X + Y < 3)$.

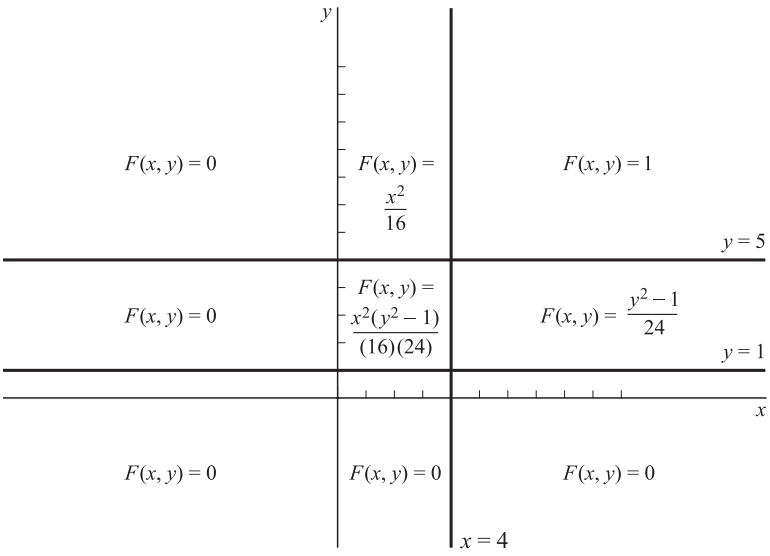


Figura 2-9

En la figura 2-10 se indica la región cuadrada $0 < x < 4$, $1 < y < 5$ en la que la función de densidad conjunta de X y Y es distinta de cero. La probabilidad buscada está dada por

$$P(X + Y < 3) = \iint_{\mathcal{R}} f(x, y)\, dx\, dy$$

donde $\mathcal{R}$ es la parte del cuadrado en la que $x + y < 3$ y se muestra sombreada en la figura 2-10. Como $f(x, y) = xy/96$ en $\mathcal{R}$, esta probabilidad está dada por

$$\int_{x=0}^{2}\int_{y=1}^{3-x}\frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=0}^{2}\left[\int_{y=1}^{3-x} xy\,dy\right]dx$$

$$= \frac{1}{96}\int_{x=0}^{2}\frac{xy^2}{2}\bigg|_{y=1}^{3-x}dx = \frac{1}{192}\int_{x=0}^{2}[x(3-x)^2 - x] = \frac{1}{48}$$

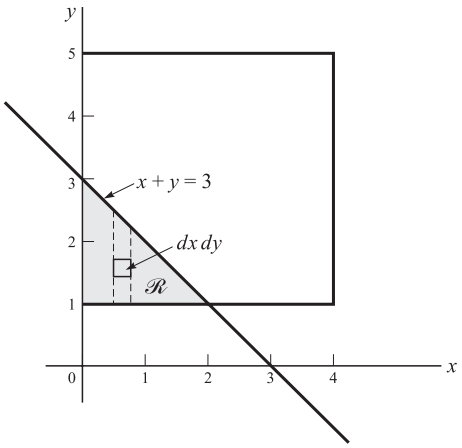


Figura 2-10

CAMBIO DE VARIABLES

2.15. Demostrar el teorema 2-1 de la página 42.

La función de probabilidad de U está dada por

$$g(u) = P(U = u) = P[\phi(X) = u] = P[X = \psi(u)] = f[\psi(u)]$$

El teorema 2-2 de la página 42 puede demostrarse de manera similar.

2.16. Demostrar el teorema 2-3 de la página 42.

Consideremos primero el caso en el que $u = \phi(x)$ o bien $x = \psi(u)$ sea una función creciente, es decir, u crece a medida que x crece (figura 2-11). En este caso, como se ve en la figura, se tiene

(1) $$P(u_1 < U < u_2) = P(x_1 < X < x_2)$$

o bien

(2) $$\int_{u_1}^{u_2} g(u)\,du = \int_{x_1}^{x_2} f(x)\,dx$$

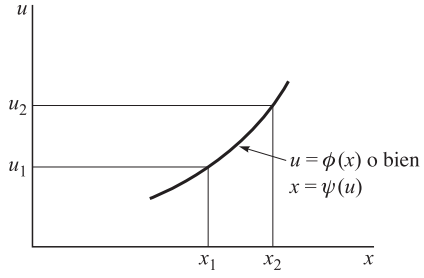


Figura 2-11

Haciendo $x = \psi(u)$ en la integral de la derecha, (2) la podemos escribir como

$$\int_{u_1}^{u_2} g(u)\,du = \int_{u_1}^{u_2} f[\psi(u)]\,\psi'(u)\,du$$

Esto es válido para toda u_1 y u_2 sólo si los integrandos son idénticos, es decir,

$$g(u) = f[\psi(u)]\psi'(u)$$

Éste es un caso especial de (34), página 42, donde $\psi'(u) > 0$ (es decir, la pendiente es positiva). Puede demostrarse que también en el caso en el que $\psi'(u) \leq 0$, es decir, en el caso en que u es función decreciente de x, (34) es válida (vea el problema 2.67). El teorema también puede demostrarse si $\psi'(u) \geq 0$ o bien $\psi'(u) < 0$.

2.17. Demostrar el teorema 2-4 de la página 42.

Primero supondremos que cuando x y y crecen, también u y v lo hacen. Como en el problema 2.16 se puede demostrar que

$$P(u_1 < U < u_2,\, v_1 < V < v_2) = P(x_1 < X < x_2,\, y_1 < Y < y_2)$$

o bien

$$\int_{v_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{x_1}^{x_2}\int_{y_1}^{y_2} f(x, y)\,dx\,dy$$

Haciendo en la integral de la derecha $x = \psi_1(u, v)$, $y = \psi_2(u, v)$, tenemos, de acuerdo con un teorema del cálculo avanzado, que

$$\int_{v_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{u_1}^{u_2}\int_{v_1}^{v_2} f[\psi_1(u, v), \psi_2(u, v)]J\,du\,dv$$

donde

$$J = \frac{\partial(x, y)}{\partial(u, v)}$$

es el *jacobiano*. En consecuencia,

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)]J$$

que es (36) de la página 42, en el caso en el que $J > 0$. De manera similar puede demostrarse (36) en el caso en el que $J < 0$.

2.18. La función de probabilidad de una variable aleatoria X es

$$f(x) = \begin{cases} 2^{-x} & x = 1, 2, 3, \ldots \\ 0 & \text{si no es así} \end{cases}$$

Encontrar la función de probabilidad de la variable aleatoria $U = X^4 + 1$.

Como $U = X^4 + 1$, la relación entre los valores u y x de las variables aleatorias U y X está dada por $u = x^4 + 1$ o bien $x = \sqrt[4]{u - 1}$, donde $u = 2, 17, 82, \ldots$ y se toma la raíz real positiva. Entonces, la función de probabilidad U que buscamos está dada por

$$g(u) = \begin{cases} 2^{-\sqrt[4]{u-1}} & u = 2, 17, 82, \ldots \\ 0 & \text{si no es así} \end{cases}$$

empleando el teorema 2-1 de la página 42 o bien el problema 2.15.

2.19. La función de probabilidad de una variable aleatoria X está dada por

$$f(x) = \begin{cases} x^2/81 & -3 < x < 6 \\ 0 & \text{si no es así} \end{cases}$$

Encontrar la densidad de probabilidad de la variable aleatoria $U = \frac{1}{3}(12 - X)$.

Tenemos $u = \frac{1}{3}(12 - x)$ o bien $x = 12 - 3u$. Por tanto, para cada valor de x hay uno y sólo un valor de u, e inversamente. Los valores de u que corresponden a $x = -3$ y $x = 6$ son $u = 5$ y $u = 2$, respectivamente. Como $\psi'(u) = dx/du = -3$, se concluye de acuerdo con el teorema 2-3 de la página 42, o con el problema 2.16, que la función de densidad de U es

$$g(u) = \begin{cases} (12 - 3u)^2/27 & 2 < u < 5 \\ 0 & \text{si no es así} \end{cases}$$

Verificación:
$$\int_2^5 \frac{(12 - 3u)^2}{27}\, du = -\frac{(12 - 3u)^3}{243}\bigg|_2^5 = 1$$

2.20. Encontrar la densidad de probabilidad de la variable aleatoria $U = X^2$, donde X es la variable aleatoria del problema 2.19.

Tenemos $u = x^2$ o bien $x = \pm\sqrt{u}$. Por tanto, a cada valor de x le corresponde un valor de u y sólo uno, pero a cada valor de $u \neq 0$ le corresponden *dos* valores de x. A los valores de x tales que $-3 < x < 6$ le corresponden los valores de u tales que $0 \leq u < 36$, como se muestra en la figura 2-12.

Como vemos en la figura, el intervalo $-3 < x \leq 3$ corresponde a $0 \leq u \leq 9$, mientras que el intervalo $3 < x < 6$ corresponde a $9 < u < 36$. En este caso no puede usarse directamente el teorema 2-3, pero se puede hacer lo siguiente. La función de distribución de U es

$$G(u) = P(U \leq u)$$

Ahora, si $0 \leq u \leq 9$, tenemos

$$G(u) = P(U \leq u) = P(X^2 \leq u) = P(-\sqrt{u} \leq X \leq \sqrt{u})$$

$$= \int_{-\sqrt{u}}^{\sqrt{u}} f(x)\, dx$$

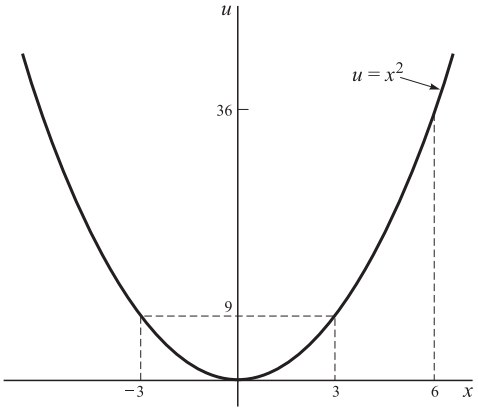


Figura 2-12

Pero si $9 < u < 36$, tenemos

$$G(u) = P(U \leq u) = P(-3 < X < \sqrt{u}) = \int_{-3}^{\sqrt{u}} f(x)\, dx$$

Dado que la función de densidad de $g(u)$ es la derivada de $G(u)$, tenemos, utilizando (12),

$$g(u) = \begin{cases} \dfrac{f(\sqrt{u}) + f(-\sqrt{u})}{2\sqrt{u}} & 0 \le u \le 9 \\ \dfrac{f(\sqrt{u})}{2\sqrt{u}} & 9 < u < 36 \\ 0 & \text{si no es así} \end{cases}$$

Utilizando la definición dada de $f(x)$, resulta:

$$g(u) = \begin{cases} \sqrt{u}/81 & 0 \le u \le 9 \\ \sqrt{u}/162 & 9 < u < 36 \\ 0 & \text{si no es así} \end{cases}$$

Verificación:

$$\int_0^9 \frac{\sqrt{u}}{81}\,du + \int_9^{36} \frac{\sqrt{u}}{162}\,du = \frac{2u^{3/2}}{243}\bigg|_0^9 + \frac{u^{3/2}}{243}\bigg|_9^{36} = 1$$

2.21. Si las variables aleatorias X y Y tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{si no es así} \end{cases}$$

(vea el problema 2.11), encontrar la función de densidad $U = X + 2Y$.

Método 1

Sean $u = x + 2y$, $v = x$, mientras que la segunda relación se elige arbitrariamente. Mediante la solución simultánea de estas dos ecuaciones obtenemos $x = v$, $y = \frac{1}{2}(u - v)$. Por tanto, la región $0 < x < 4$, $1 < y < 5$ corresponde a la región $0 < v < 4$, $2 < u - v < 10$ que se muestra sombreada en la figura 2-13.

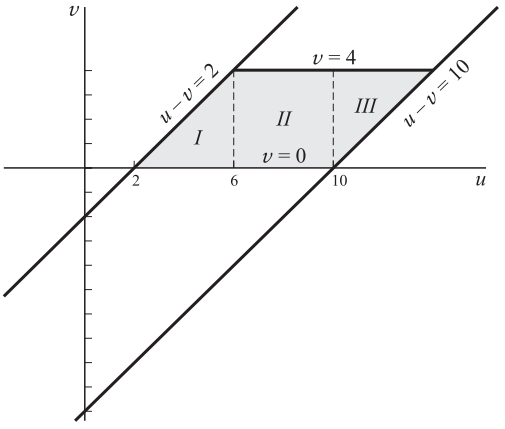


Figura 2-13

El jacobiano está dado por

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

Entonces, de acuerdo con el teorema 2-4, la función de densidad conjunta de U y V es

$$g(u, v) = \begin{cases} v(u - v)/384 & 2 < u - v < 10,\ 0 < v < 4 \\ 0 & \text{si no es así} \end{cases}$$

La función de densidad marginal de U está dada por

$$g_1(u) = \begin{cases} \displaystyle\int_{v=0}^{u-2} \frac{v(u - v)}{384}\, dv & 2 < u < 6 \\ \displaystyle\int_{v=0}^{4} \frac{v(u - v)}{384}\, dv & 6 < u < 10 \\ \displaystyle\int_{v=u-10}^{4} \frac{v(u - v)}{384}\, dv & 10 < u < 14 \\ 0 & \text{si no es así} \end{cases}$$

como puede verse observando las regiones sombreadas *I*, *II* y *III* de la figura 2-13. Cuando se realiza la integración encontramos

$$g_1(u) = \begin{cases} (u - 2)^2(u + 4)/2\,304 & 2 < u < 6 \\ (3u - 8)/144 & 6 < u < 10 \\ (348u - u^3 - 2\,128)/2\,304 & 10 < u < 14 \\ 0 & \text{si no es así} \end{cases}$$

Este resultado se puede verificar si se demuestra que la integral de $g_1(u)$ es igual a 1.

Método 2
La función de distribución de la variable aleatoria $X + 2Y$ está dada por

$$P(X + 2Y \leq u) = \iint_{x+2y\leq u} f(x, y)\, dx\, dy = \iint_{\substack{x+2y\leq u \\ 0<x<4 \\ 1<y<5}} \frac{xy}{96}\, dx\, dy$$

Para $2 < u < 6$, observando la figura 2-14 puede verse que la última integral es igual a

$$\int_{x=0}^{u-2}\int_{y=1}^{(u-x)/2} \frac{xy}{96}\, dx\, dy = \int_{x=0}^{u-2}\left[\frac{x(u - x)^2}{768} - \frac{x}{192}\right] dx$$

La derivada con respecto a u es $(u - 2)^2(u + 4)2\,304$. De manera similar puede obtenerse el resultado del método 1 para $6 < u < 10$, etcétera.

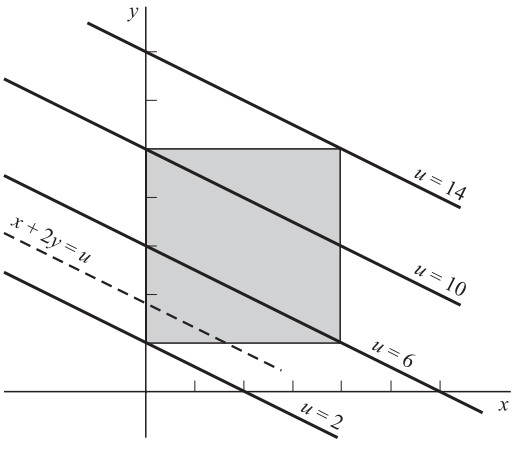


Figura 2-14

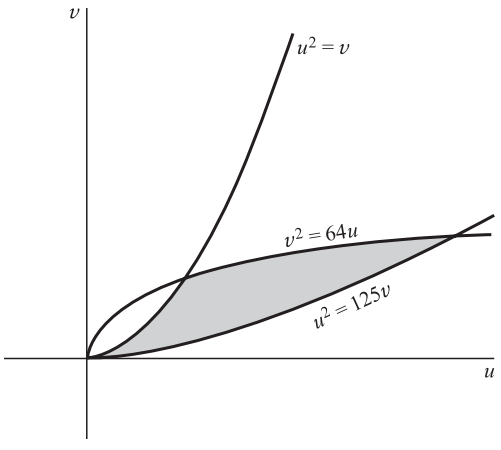


Figura 2-15

2.22. Si las variables aleatorias X y Y tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{si no es así} \end{cases}$$

(vea el problema 2.11), encuentre la función de densidad conjunta de $U = XY^2$, $V = X^2Y$.

Consideremos $u = xy^2$, $v = x^2y$. Dividiendo estas ecuaciones se obtiene $x/y = u/v$, de manera que $y = ux/v$. Esto lleva a la solución simultánea $x = v^{2/3}u^{-1/3}$, $y = u^{2/3}v^{-1/3}$. La imagen de $0 < x < 4$, $1 < y < 5$ en el plano uv está dada por

$$0 < v^{2/3}u^{-1/3} < 4 \qquad 1 < u^{2/3}v^{-1/3} < 5$$

que es equivalente a

$$v^2 < 64u \qquad v < u^2 < 125v$$

Esta región es la que se muestra sombreada en la figura 2-15.

El jacobiano está dado por

$$J = \begin{vmatrix} -\frac{1}{3}v^{2/3}u^{-4/3} & \frac{2}{3}v^{-1/3}u^{-1/3} \\ \frac{2}{3}u^{-1/3}v^{-1/3} & -\frac{1}{3}u^{2/3}v^{-4/3} \end{vmatrix} = -\frac{1}{3}u^{-2/3}v^{-2/3}$$

Por tanto, la función de densidad conjunta de U y V es, de acuerdo con el teorema 2-4,

$$g(u, v) = \begin{cases} \dfrac{(v^{2/3}u^{-1/3})(u^{2/3}v^{-1/3})}{96}\left(\frac{1}{3}u^{-2/3}v^{-2/3}\right) & v^2 < 64u, v < u^2 < 125v \\ 0 & \text{si no es así} \end{cases}$$

o bien

$$g(u, v) = \begin{cases} u^{-1/3}v^{-1/3}/288 & v^2 < 64u, \quad v < u^2 < 125v \\ 0 & \text{si no es así} \end{cases}$$

CONVOLUCIONES

2.23. Sean X y Y variables aleatorias con función de densidad conjunta $f(x, y)$. Demostrar que la función de densidad de $U = X + Y$ es

$$g(u) = \int_{-\infty}^{\infty} f(v, u - v)dv$$

Método 1

Sean $U = X + Y$, $V = X$, donde la segunda ecuación se ha agregado de manera arbitraria. A cada una de estas ecuaciones corresponden $u = x + y$, $v = x$ o bien $x = v$, $y = u - v$. El jacobiano de esta transformación está dado por

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Por tanto, de acuerdo con el teorema 2-4, de la página 42, la función de densidad conjunta de U y V es

$$g(u, v) = f(v, u - v)$$

De (26), página 41, se deduce que la función de densidad marginal de U es

$$g(u) = \int_{-\infty}^{\infty} f(v, u - v)dv$$

Método 2

La función de distribución de $U = X + Y$ es igual a la integral doble de $f(x, y)$ sobre la región definida por $x + y \leq u$, es decir,

$$G(u) = \iint_{x+y\leq u} f(x, y)\, dx\, dy$$

Dado que esta región se encuentra debajo de la recta $x + y = u$, como se indica mediante la región sombreada de la figura 2-16, vemos que

$$G(u) = \int_{x=-\infty}^{\infty} \left[\int_{y=-\infty}^{u-x} f(x, y)\, dy \right] dx$$

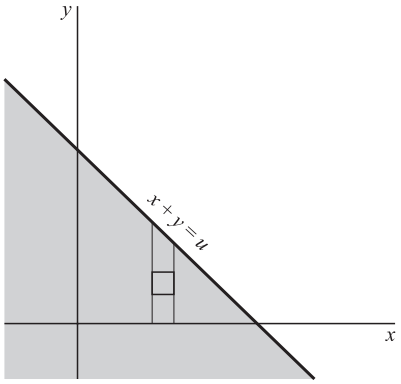


Figura 2-16

La función de densidad de U es la derivada de $G(u)$ respecto de u y está dada por

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx$$

usando (12) primero sobre la integral de x y después sobre la integral de y.

2.24. Repetir el problema 2.23 para el caso en el que X y Y sean variables aleatorias independientes con funciones de densidad $f_1(x), f_2(y)$, respectivamente.

En este caso, la función de densidad conjunta es $f(x, y) = f_1(x)f_2(y)$, de manera que de acuerdo con el problema 2.23, la función de densidad de $U = X + Y$ es

$$g(u) = \int_{-\infty}^{\infty} f_1(v)f_2(u - v)dv = f_1 * f_2$$

que es la *convolución* de f_1 y f_2.

2.25. Si X y Y son variables aleatorias independientes con funciones de densidad

$$f_1(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad f_2(y) = \begin{cases} 3e^{-3y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

encontrar la función de densidad de su suma, $U = X + Y$.

De acuerdo con el problema 2.24, la función de densidad que buscamos es la convolución de f_1 y f_2 y está dada por

$$g(u) = f_1 * f_2 = \int_{-\infty}^{\infty} f_1(v)f_2(u - v)\, dv$$

En este integrando, f_1 desaparece cuando $v < 0$ y f_2 es cero cuando $v > u$. Por tanto,

$$g(u) = \int_0^u (2e^{-2v})(3e^{-3(u-v)})\, dv$$

$$= 6e^{-3u} \int_0^u e^v dv = 6e^{-3u}(e^u - 1) = 6(e^{-2u} - e^{3u})$$

si $u \geq 0$ y $g(u) = 0$ si $u < 0$.

Verificación: $$\int_{-\infty}^{\infty} g(u)\,du = 6\int_{0}^{\infty}(e^{-2u} - e^{-3u})\,du = 6\left(\frac{1}{2} - \frac{1}{3}\right) = 1$$

2.26. Demostrar que $f_1 * f_2 = f_2 * f_1$ (propiedad 1, página 43).

Tenemos

$$f_1 * f_2 = \int_{v=-\infty}^{\infty} f_1(v) f_2(u - v)\,dv$$

Haciendo $w = u - v$ de manera que $v = u - w$, $dv = -dw$, obtenemos

$$f_1 * f_2 = \int_{w=\infty}^{-\infty} f_1(u - w) f_2(w)(-dw) = \int_{w=-\infty}^{\infty} f_2(w) f_1(u - w)\,dw = f_2 * f_1$$

DISTRIBUCIONES CONDICIONALES

2.27. Dada la distribución del problema 2.8, encontrar *a)* $f(y\,|\,2)$, *b)* $P(Y = 1\,|\,X = 2)$.

a) Usando los resultados de los problemas 2.8 y 2.9, tenemos

$$f(y\,|\,x) = \frac{f(x, y)}{f_1(x)} = \frac{(2x + y)/42}{f_1(x)}$$

de manera que si $x = 2$

$$f(y\,|\,2) = \frac{(4 + y)/42}{11/21} = \frac{4 + y}{22}$$

b) $$P(Y = 1\,|\,X = 2) = f(1\,|\,2) = \frac{5}{22}$$

2.28. Si X y Y tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} \frac{3}{4} + xy & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{si no es así} \end{cases}$$

encontrar *a)* $f(y\,|\,x)$, *b)* $P(Y > \frac{1}{2}\,|\,\frac{1}{2} < X < \frac{1}{2} + dx)$.

a) Para $0 < x < 1$,

$$f_1(x) = \int_0^1 \left(\frac{3}{4} + xy\right) dy = \frac{3}{4} + \frac{x}{2}$$

y $$f(y\,|\,x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} \dfrac{3 + 4xy}{3 + 2x} & 0 < y < 1 \\ 0 & \text{otros valores de } y \end{cases}$$

Para otros valores de x, $f(y\,|\,x)$ no está definida.

b) $$P(Y > \tfrac{1}{2}\,|\,\tfrac{1}{2} < X < \tfrac{1}{2} + dx) = \int_{1/2}^{\infty} f(y\,|\,\tfrac{1}{2})\,dy = \int_{1/2}^{1} \frac{3 + 2y}{4}\,dy = \frac{9}{16}$$

2.29. La función de densidad conjunta de las variables aleatorias X y Y está dada por

$$f(x, y) = \begin{cases} 8xy & 0 \leq x \leq 1,\ 0 \leq y \leq x \\ 0 & \text{si no es así} \end{cases}$$

Encontrar *a*) la densidad marginal de X, *b*) la densidad marginal de Y, *c*) la densidad condicional de X, *d*) la densidad condicional de Y.

En la figura 2-17 se muestra la región en la que $f(x, y)$ es distinta de cero.

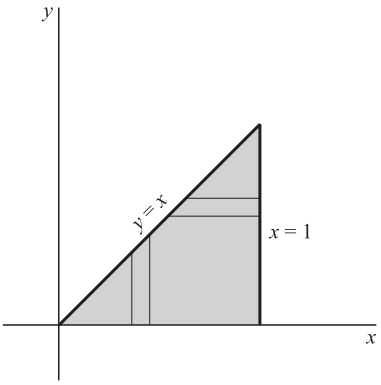


Figura 2-17

a) Para obtener la densidad marginal de X, se fija x y se integra respecto de y desde 0 hasta x, como se muestra mediante la franja vertical de la figura 2-17. El resultado es

$$f_1(x) = \int_{y=0}^{x} 8xy\,dy = 4x^3$$

para $0 < x < 1$. Para todos los demás valores de $x, f_1(x) = 0$.

b) De manera similar, la densidad marginal de Y se obtiene fijando y e integrando respecto a x desde $x = y$ hasta $x = 1$ como se indica mediante la franja horizontal de la figura 2-17. El resultado, para $0 < y < 1$, es

$$f_2(y) = \int_{x=y}^{1} 8xy\,dx = 4y(1 - y^2)$$

Para todos los demás valores de $y, f_2(y) = 0$.

c) La función de densidad condicional de X es, para $0 < y < 1$,

$$f_1(x|y) = \frac{f(x, y)}{f_2(y)} = \begin{cases} 2x/(1 - y^2) & y \le x \le 1 \\ 0 & \text{otros valores de } x \end{cases}$$

La función de densidad condicional no está definida cuando $f_2(y) = 0$.

d) La función de densidad condicional de Y es, para $0 < x < 1$,

$$f_2(y|x) = \frac{f(x, y)}{f_1(x)} = \begin{cases} 2y/x^2 & 0 \le y \le x \\ 0 & \text{otros valores de } y \end{cases}$$

La función de densidad condicional no está definida cuando $f_1(x) = 0$.

Verificación: $\displaystyle\int_0^1 f_1(x)\,dx = \int_0^1 4x^3\,dx = 1, \quad \int_0^1 f_2(y)\,dy = \int_0^1 4y(1 - y^2)\,dy = 1$

$$\int_y^1 f_1(x|y)\,dx = \int_y^1 \frac{2x}{1 - y^2}\,dx = 1$$

$$\int_0^x f_2(y|x)\,dy = \int_0^x \frac{2y}{x^2}\,dy = 1$$

2.30. Determinar si las variables aleatorias del problema 2.29 son independientes.

En la región sombreada de la figura 2-17, $f(x, y) = 8xy$, $f_1(x) = 4x^3$, $f_2(y) = 4y\,(1 - y^2)$. Por tanto, $f(x, y) \neq f_1(x)\,f_2(y)$ y X y Y son dependientes.

Hay que hacer notar que de $f(x, y) = 8xy$ no obtenemos que $f(x, y)$ pueda expresarse como una función sólo de x multiplicada por una función sólo de y. Esto se debe a la presencia de la restricción $0 \le y \le x$. Si esta restricción se sustituyera por alguna otra restricción para y que no dependiera de x (como en el problema 2.21), tal conclusión sería válida.

APLICACIONES DEL CÁLCULO DE LA PROBABILIDAD GEOMÉTRICA

2.31. Una persona que juega dardos determina que la probabilidad de que el dardo caiga entre r y $r + dr$ es

$$P(r \leq R \leq r + dr) = c\left[1 - \left(\frac{r}{a}\right)^2\right]dr$$

En este caso, R es la distancia al centro del blanco, c es una constante y a es el radio del blanco (vea la figura 2-18). Encontrar la probabilidad de dar en el blanco, que se supone tiene radio b. Supondremos que siempre se da en él.

La función de densidad está dada por

$$f(r) = c\left[1 - \left(\frac{r}{a}\right)^2\right]$$

Como siempre se da en el blanco, tenemos

$$c\int_0^a \left[1 - \left(\frac{r}{a}\right)^2\right]dr = 1$$

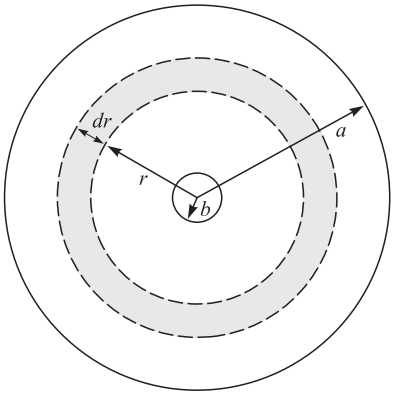


Figura 2-18

de donde $c = 3/2a$. Entonces, la probabilidad de dar en el blanco es

$$\int_0^b f(r)\,dr = \frac{3}{2a}\int_0^b \left[1 - \left(\frac{r}{a}\right)^2\right]dr = \frac{b(3a^2 - b^2)}{2a^3}$$

2.32. En el intervalo $0 \leq x \leq 1$ se eligen dos puntos al azar. Determinar la probabilidad de que la suma de sus cuadrados sea menor que 1.

Sean X y Y las variables aleatorias relacionadas con los puntos dados. Puesto que suponemos que intervalos iguales tienen las mismas probabilidades, las funciones de densidad de X y de Y están dadas, respectivamente, por

$$(1) \qquad f_1(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{si no es así} \end{cases} \qquad f_2(y) = \begin{cases} 1 & 0 \leq y \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

Entonces, dado que X y Y son independientes, la función de densidad conjunta está dada por

$$(2) \qquad f(x, y) = f_1(x)f_2(y) = \begin{cases} 1 & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

Se concluye que la probabilidad que buscamos está dada por

$$(3) \qquad P(X^2 + Y^2 \leq 1) = \iint_{\mathcal{R}} dx\,dy$$

donde $\mathcal{R}$ es la región definida por $x^2 + y^2 \leq 1, x \geq 0, y \geq 0$, que es un cuarto de un círculo de radio 1 (figura 2-19). Ahora como (3) representa el área de $\mathcal{R}$, se ve que la probabilidad que buscamos es $\pi/4$.

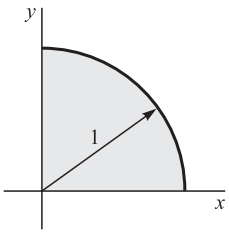


Figura 2-19

PROBLEMAS DIVERSOS

2.33. Se supone que las variables aleatorias X y Y tienen una función de densidad conjunta dada por

$$f(x, y) = \begin{cases} c(2x + y) & 2 < x < 6, 0 < y < 5 \\ 0 & \text{si no es así} \end{cases}$$

Determinar *a*) la constante c, *b*) las funciones de distribución marginal de X y de Y, *c*) las funciones de densidad marginal de X y de Y, *d*) $P(3 < X < 4, Y > 2)$, *e*) $P(X > 3)$, *f*) $P(X + Y > 4)$, *g*) la función de distribución conjunta, *h*) si X y Y son independientes.

a) La probabilidad total está dada por

$$\int_{x=2}^{6}\int_{y=0}^{5} c(2x + y)\,dx\,dy = \int_{x=2}^{6} c\left(2xy + \frac{y^2}{2}\right)\Bigg|_0^5 dx$$

$$= \int_{x=2}^{6} c\left(10x + \frac{25}{2}\right)dx = 210c$$

Para que esto sea igual a 1, debemos tener $c = 1/210$.

b) La función de distribución marginal de X es

$$F_1(x) = P(X \le x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} 0\,du\,dv = 0 & x < 2 \\ \displaystyle\int_{u=2}^{x}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = \frac{2x^2 + 5x - 18}{84} & 2 \le x < 6 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = 1 & x \ge 6 \end{cases}$$

La función de distribución marginal de Y es

$$F_2(y) = P(Y \le y) = \int_{u=-\infty}^{\infty}\int_{v=-\infty}^{y} f(u, v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{\infty}\int_{v=-8}^{y} 0\,du\,dv = 0 & y < 0 \\ \displaystyle\int_{u=0}^{6}\int_{v=0}^{y} \frac{2u + v}{210}\,du\,dv = \frac{y^2 + 16y}{105} & 0 \le y < 5 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u + v}{210}\,du\,dv = 1 & y \ge 5 \end{cases}$$

c) De acuerdo con el inciso *b*), la función de densidad marginal de X es

$$f_1(x) = \frac{d}{dx}F_1(x) = \begin{cases} (4x + 5)/84 & 2 < x < 6 \\ 0 & \text{si no es así} \end{cases}$$

De acuerdo con el inciso *b*), la función de densidad marginal de Y es

$$f_2(y) = \frac{d}{dy}F_2(y) = \begin{cases} (2y + 16)/105 & 0 < y < 5 \\ 0 & \text{si no es así} \end{cases}$$

d)
$$P(3 < X < 4, Y > 2) = \frac{1}{210}\int_{x=3}^{4}\int_{y=2}^{5}(2x + y)\,dx\,dy = \frac{3}{20}$$

e)
$$P(X > 3) = \frac{1}{210}\int_{x=3}^{6}\int_{y=0}^{5}(2x + y)\,dx\,dy = \frac{23}{28}$$

f)
$$P(X + Y > 4) = \iint_{\mathcal{R}} f(x, y)\,dx\,dy$$

donde $\mathcal{R}$ es la región sombreada que se muestra en la figura 2-20. Aunque esto puede obtenerse, es más fácil usar el hecho de que

$$P(X + Y > 4) = 1 - P(X + Y \leq 4) = 1 - \iint_{\mathcal{R}'} f(x, y)\,dx\,dy$$

donde $\mathcal{R}'$ es la región doblemente sombreada de la figura 2-20. Tenemos

$$P(X + Y \leq 4) = \frac{1}{210}\int_{x=2}^{4}\int_{y=0}^{4-x}(2x + y)\,dx\,dy = \frac{2}{35}$$

Por tanto, $P(X + Y > 4) = 33/35$.

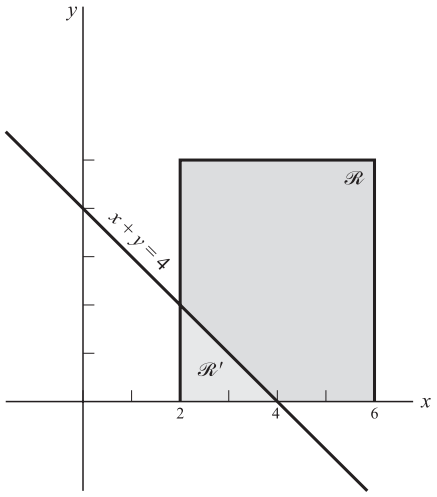


Figura 2-20

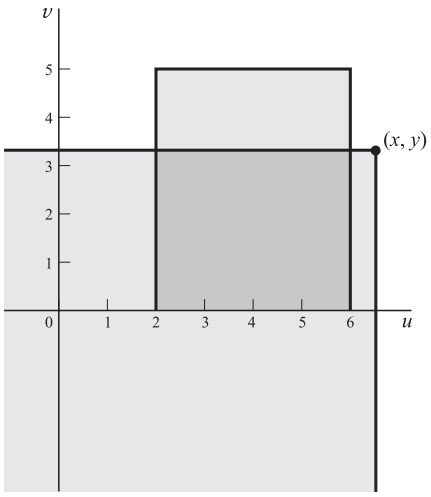


Figura 2-21

g) La función de distribución conjunta es

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{y} f(u, v)\,du\,dv$$

En el plano uv (figura 2-21) la región de integración es la intersección del plano $u \leq x$, $v \leq y$, además, el rectángulo $2 < u < 6$, $0 < v < 5$ [en donde $f(u, v)$ es distinta de cero]. Para (x, y) ubicados como en la figura, se tiene

$$F(x, y) = \int_{u=2}^{6}\int_{v=0}^{y}\frac{2u + v}{210}\,du\,dv = \frac{16y + y^2}{105}$$

Cuando (x, y) se encuentra dentro del rectángulo, obtenemos otra expresión, etc. En la figura 2-22 se muestran todos los resultados.

h) Estas variables aleatorias son dependientes ya que

$$f(x, y) \neq f_1(x)f_2(y)$$

o, lo que es equivalente, $F(x, y) \neq F_1(x)F_2(y)$.

2.34. Considere la función de densidad de X siguiente:

$$f(x) = \begin{cases} 6x(1 - x) & 0 < x < 1 \\ 0 & \text{si no es así} \end{cases}$$

Encontrar una función $Y = h(X)$ que tenga la función de densidad

$$g(y) = \begin{cases} 12y^3(1 - y^2) & 0 < y < 1 \\ 0 & \text{si no es así} \end{cases}$$

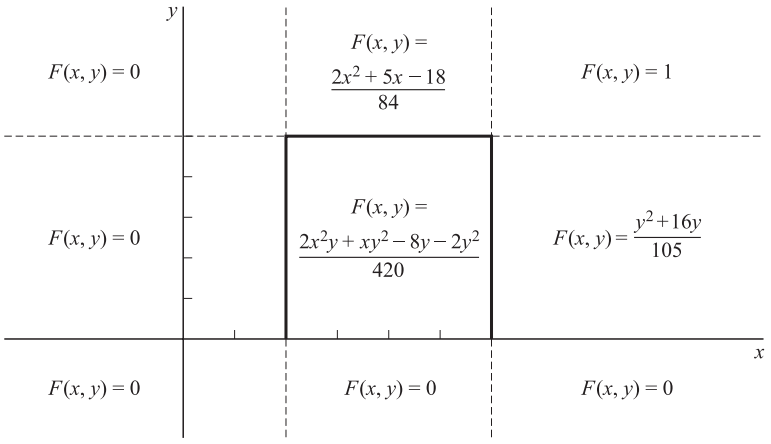


Figura 2-22

Suponemos que la función desconocida h es una función tal que los intervalos $X \leq x$ y $Y \leq y + h(x)$ se corresponden uno a uno, de manera continua. En consecuencia, $P(X \leq x) = P(Y \leq y)$, es decir, las funciones de distribución de X y de Y deben ser iguales. Por tanto, para $0 < x, y < 1$,

$$\int_0^x 6u(1 - u)\,du = \int_0^y 12v^3(1 - v^2)\,dv$$

o bien

$$3x^2 - 2x^3 = 3y^4 - 2y^6$$

Por inspección, una solución es $x = y^2$ o bien $y = h(x) = +\sqrt{x}$, solución que tiene las propiedades deseadas. Por tanto, $Y = +\sqrt{X}$.

2.35. Encontrar la función de densidad de $U = XY$ si la función de densidad conjunta de X y Y es $f(x, y)$.

Método 1

Sea $U = XY$ y $V = X$, que corresponden a $u = xy$, $v = x$ o bien $x = v$, y $= u/v$. Entonces, el jacobiano está dado por

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ v^{-1} & -uv^{-2} \end{vmatrix} = -v^{-1}$$

Por tanto, la función de densidad conjunta de U y V es

$$g(u, v) = \frac{1}{|v|} f\left(v, \frac{u}{v}\right)$$

de donde la función de densidad marginal de U es

$$g(u) = \int_{-\infty}^{\infty} g(u, v)\, dv = \int_{-\infty}^{\infty} \frac{1}{|v|} f\left(v, \frac{u}{v}\right) dv$$

Método 2
La función de distribución de U es

$$G(u) = \iint_{xy \le u} f(x, y)\, dx\, dy$$

En la figura 2-23 se muestra sombreada la región de integración para $u \ge 0$. Vemos que

$$G(u) = \int_{-\infty}^{0} \left[\int_{u/x}^{\infty} f(x, y)\, dy \right] dx + \int_{0}^{\infty} \left[\int_{-\infty}^{u/x} f(x, y)\, dy \right] dx$$

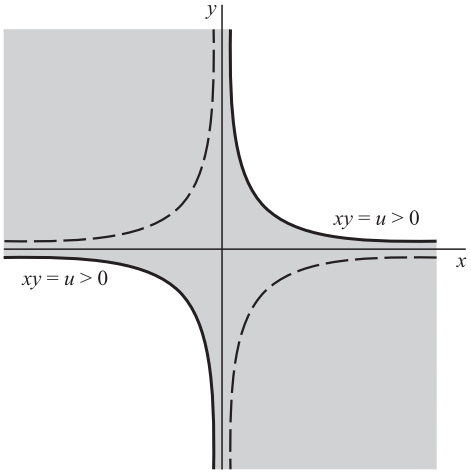


Figura 2-23

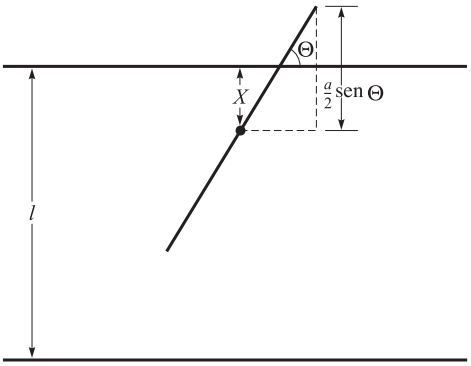


Figura 2-24

Diferenciando respecto a u obtenemos

$$g(u) = \int_{-\infty}^{0} \left(\frac{-1}{x}\right) f\left(x, \frac{u}{x}\right) dx + \int_{0}^{\infty} \frac{1}{x} f\left(x, \frac{u}{x}\right) dx = \int_{-\infty}^{\infty} \frac{1}{|x|} f\left(x, \frac{u}{x}\right) dx$$

Para $u < 0$ obtenemos el mismo resultado, cuando la región de integración está limitada por la hipérbola punteada de la figura 2-23.

2.36. Un piso tiene rectas paralelas que equidistan una distancia l. Sobre el piso se deja caer al azar una aguja de longitud $a < l$. Encontrar la probabilidad de que la aguja cruce una de las rectas. (Este problema se conoce como el *problema de la aguja de Buffon*.)

Sea X una variable aleatoria que dé la distancia del punto medio de la aguja a la recta más cercana (figura 2-24). Sea Θ una variable aleatoria que dé el ángulo agudo entre la aguja (o su extensión) y la recta. Los valores particulares de X y θ se denotan como X y Θ. Se ve que X puede tomar cualquier valor entre 0 y $l/2$, de manera que $0 \le x \le l/2$. Por otro lado, Θ puede tomar cualquier valor entre 0 y $\pi/2$. Se infiere que

$$P(x < X \le x + dx) = \frac{2}{l}\, dx \qquad P(\theta \le \Theta + d\theta) = \frac{2}{\pi}\, d\theta$$

es decir, las funciones de densidad de X y Θ están dadas por $f_1(x) = 2/l, f_2(\theta) = 2/\pi$. Para verificar, observemos que

$$\int_0^{l/2} \frac{2}{l}\, dx = 1 \quad \int_0^{\pi/2} \frac{2}{\pi}\, d\theta = 1$$

Como X y Θ son independientes, la función de densidad conjunta es

$$f(x, \theta) = \frac{2}{l} \cdot \frac{2}{\pi} = \frac{4}{l\pi}$$

En la figura 2-24 vemos que la aguja toca una recta cuando $X \leq (a/2)$ sen Θ. La probabilidad de este evento está dada por

$$\frac{4}{l\pi} \int_{\theta=0}^{\pi/2} \int_{x=0}^{(a/2)\,\text{sen}\,\theta} dx\, d\theta = \frac{2a}{l\pi}$$

Cuando la expresión anterior se iguala a la frecuencia de caídas observadas en forma experimental, se obtienen valores exactos de π, lo que indica que el modelo de probabilidad descrito es apropiado.

2.37. Dos personas convienen en encontrarse entre las 2:00 p.m. y las 3:00 p.m. También acuerdan en que cada una esperará no más de 15 minutos a la otra. ¿Cuál es la probabilidad de que se encuentren?

Sean X y Y las variables aleatorias que representan el tiempo de llegada de cada una de estas dos personas, medido en fracciones de hora después de las 2:00 p.m. Suponiendo que intervalos iguales de tiempo tengan probabilidades iguales de la llegada de estas personas, las funciones de densidad de X y Y están dadas respectivamente por

$$f_1(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

$$f_2(y) = \begin{cases} 1 & 0 \leq y \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

En consecuencia, como X y Y son independientes, la función de densidad conjunta es

$$f(x, y) = f_1(x) f_2(y) = \begin{cases} 1 & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{si no es así} \end{cases} \tag{1}$$

Dado que 15 minutos $= \frac{1}{4}$ de hora, la probabilidad que buscamos es

$$P\left(|X - Y| \leq \frac{1}{4}\right) = \iint_{\mathcal{R}} dx\, dy \tag{2}$$

donde $\mathcal{R}$ en la región que se muestra sombreada en la figura 2-25. El lado derecho de (2) es el área de esta región, que es igual a $1 - (\frac{3}{4})(\frac{3}{4}) = \frac{7}{16}$, dado que el cuadrado tiene área 1, mientras que los dos triángulos de las esquinas tienen un área de $\frac{1}{2}(\frac{3}{4})(\frac{3}{4})$ cada uno. Por tanto, la probabilidad que se busca es $7/16$.

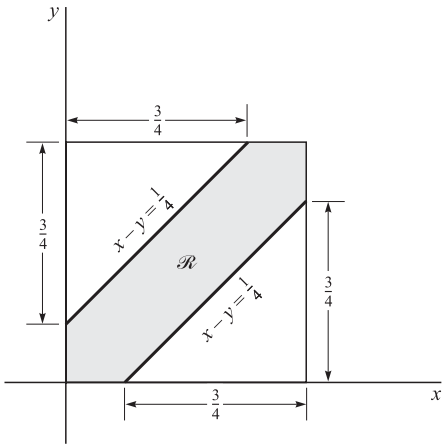


Figura 2-25

PROBLEMAS SUPLEMENTARIOS

VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD DISCRETAS

2.38. Una moneda se lanza tres veces. Si X es una variable aleatoria que da el número de las caras que se obtienen, construya una tabla que muestre la distribución de probabilidad de X.

2.39. Una urna contiene 5 canicas blancas y 3 negras. Si se extraen 2 canicas al azar sin reposición y X denota el número de canicas blancas, encuentre la distribución de probabilidad de X.

2.40. Repita el problema 2.39 si las canicas que se extraen se reponen.

2.41. Sea Z una variable aleatoria que da el número de caras menos el número de cruces en 2 lanzamientos de una moneda no cargada. Encuentre la distribución de probabilidad de Z. Compare con los resultados de los ejemplos 2.1 y 2.2.

2.42. Sea X una variable aleatoria que da el número de ases en una extracción al azar de 4 cartas de una baraja ordinaria de 52 cartas. Construya una tabla que muestre la distribución de probabilidad de X.

FUNCIONES DE DISTRIBUCIÓN DISCRETAS

2.43. En la tabla 2-7 se muestra la función de probabilidad de una variable aleatoria X. Construya una tabla que dé la función de distribución de X.

Tabla 2-7

x	1	2	3
$f(x)$	$1/2$	$1/3$	$1/6$

Tabla 2-8

x	1	2	3	4
$F(x)$	$1/8$	$3/8$	$3/4$	1

2.44. Obtenga la función de distribución correspondiente a *a*) el problema 2.38, *b*) el problema 2.39, *c*) el problema 2.40.

2.45. Obtenga la función de distribución correspondiente a *a*) el problema 2.41, *b*) el problema 2.42.

2.46. La tabla 2-8 muestra la función de distribución de una variable aleatoria X. Determine *a*) la función de probabilidad, *b*) $P(1 \leq X \leq 3)$, *c*) $P(X \geq 2)$, *d*) $P(X < 3)$, *e*) $P(X > 1.4)$.

VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD CONTINUAS

2.47. Una variable aleatoria X tiene la función de densidad

$$f(x) = \begin{cases} ce^{-3x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Encuentre *a*) la constante c, *b*) $P(l < X < 2)$, *c*) $P(X \geq 3)$, *d*) $P(X < 1)$.

2.48. Encuentre la función de distribución de la variable aleatoria del problema 2.47. Represente en forma gráfica las funciones de densidad y de distribución, y describa la relación entre ellas.

2.49. Una variable aleatoria X tiene la función de densidad

$$f(x) = \begin{cases} cx^2 & 1 \leq x \leq 2 \\ cx & 2 < x < 3 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) la constante c, *b*) $P(X > 2)$, *c*) $P(1/2 < X < 3/2)$.

2.50. Encuentre la función de distribución de la variable aleatoria X del problema 2.49.

2.51. La función de distribución de una variable aleatoria X está dada por

$$F(x) = \begin{cases} cx^3 & 0 \le x < 3 \\ 1 & x \ge 3 \\ 0 & x < 0 \end{cases}$$

Si $P(X = 3) = 0$, encuentre *a*) la constante c, *b*) la función de densidad, *c*) $P(X > 1)$, *d*) $P(1 < X < 2)$.

2.52. ¿Puede la función

$$F(x) = \begin{cases} c(1 - x^2) & 0 \le x \le 1 \\ 0 & \text{si no es así} \end{cases}$$

ser una función de distribución? Explique su respuesta.

2.53. Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} cx & 0 \le x \le 2 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) el valor de la constante c, *b*) $P(\frac{1}{2} < X < \frac{3}{2})$, *c*) $P(X > 1)$, *d*) la función de distribución.

DISTRIBUCIONES CONJUNTAS Y VARIABLES INDEPENDIENTES

2.54. La función de probabilidad conjunta de dos variables aleatorias X y Y está dada por $f(x, y) = cxy$ para $x = 1, 2, 3$ y $y = 1, 2, 3$, e igual a cero si no es así. Encuentre *a*) la constante c, *b*) $P(X = 2, Y = 3)$, *c*) $P(1 \le X \le 2, Y \le 2)$, *d*) $P(X \ge 2)$, *e*) $P(Y < 2)$, *f*) $P(X = 1)$, *g*) $P(Y = 3)$.

2.55. Encuentre las funciones de probabilidad marginal de *a*) X y *b*) Y de las variables aleatorias del problema 2.54. *c*) Determine si X y Y son independientes.

2.56. Sean X y Y variables aleatorias continuas con función de densidad conjunta

$$f(x, y) = \begin{cases} c(x^2 + y^2) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{si no es así} \end{cases}$$

Determine *a*) la constante c, *b*) $P(X < \frac{1}{2}, Y > \frac{1}{2})$, *c*) $P(\frac{1}{4} < X < \frac{3}{4})$, *d*) $P(Y < \frac{1}{2})$, *e*) si X y Y son independientes.

2.57. Encuentre las funciones de distribución marginal *a*) de X y *b*) de Y de la función de densidad del problema 2.56.

DISTRIBUCIONES CONDICIONALES Y FUNCIONES DE DENSIDAD

2.58. Encuentre la función de probabilidad condicional *a*) de X dada Y, *b*) de Y dada X en el caso de la distribución del problema 2.54.

2.59. Sea

$$f(x, y) = \begin{cases} x + y & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre la función de densidad condicional de *a*) X dado Y, *b*) Y dado X.

2.60. Encuentre la densidad condicional de *a*) X dado Y, *b*) Y dado X en el caso de la distribución del problema 2.56.

2.61. Sea

$$f(x, y) = \begin{cases} e^{-(x+y)} & x \ge 0, y \ge 0 \\ 0 & \text{si no es así} \end{cases}$$

la función de densidad conjunta de X y Y. Encuentre la función de densidad condicional de *a*) X dada Y, *b*) Y dada X.

CAMBIO DE VARIABLES

2.62. Sea la función de densidad de X

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Encuentre la función de densidad de $Y = X^2$.

2.63. *a*) Si la función de densidad de X es $f(x)$ encuentre la función de densidad de X^3. *b*) Ilustre el resultado del inciso *a*) empleando la función

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

y compruebe la respuesta.

2.64. Si X tiene la función de densidad $f(x) = 2(\pi)^{-1/2}e^{-x^2/2}$, $-\infty < x < \infty$, encuentre la función de densidad de $Y = X^2$.

2.65. Verifique que en el método 1 del problema 2.21 la integral de $g_1(u)$ es igual a 1.

2.66. Si la densidad de X es $f(x) = 1/\pi(x^2 + 1)$, $-\infty < x < \infty$, encuentre la densidad de $Y = \tan^{-1} X$.

2.67. En el método dos del problema 2.21 realice lo que haga falta para encontrar $g_1(u)$ y verifique su respuesta.

2.68. La función de densidad de X es

$$f(x) = \begin{cases} 1/2 & -1 < x < 1 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre la función de densidad de *a*) $3X - 2$, *b*) $X^3 + 1$.

2.69. Verifique la función de densidad conjunta encontrada en el problema 2.22 mediante integración directa.

2.70. La función de densidad conjunta de X y Y es

$$f(x, y) = \begin{cases} e^{-(x+y)} & x \geq 0, y \geq 0 \\ 0 & \text{si no es así} \end{cases}$$

Si $U = X/Y$, $V = X + Y$, encuentre la función de densidad conjunta de U y V.

2.71. Use el problema 2.22 para determinar la función de densidad de *a*) $U = XY^2$, *b*) $V = X^2Y$.

2.72. Sean X y Y variables aleatorias cuya función de densidad conjunta es $f(x, y) = (2\pi)^{-1}e^{-(x^2+y^2)}$, $-\infty < x < \infty$, $-\infty < y < \infty$. Si R y Θ son otras variables aleatorias tales que $X = R\cos\Theta$, $Y = R\,\text{sen}\,\Theta$, muestre que la función de densidad de R es

$$g(r) = \begin{cases} re^{-r^2/2} & r \geq 0 \\ 0 & r < 0 \end{cases}$$

2.73. Sea
$$f(x, y) = \begin{cases} 1 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{si no es así} \end{cases}$$
la función de densidad conjunta de X y Y. Encuentre la función de densidad de $Z = XY$.

CONVOLUCIONES

2.74. Sean X y Y variables aleatorias independientes distribuidas de manera idéntica con función de densidad
$$f(t) = \begin{cases} 1 & 0 \le t \le 1 \\ 0 & \text{si no es así} \end{cases}$$
Encuentre la función de densidad de $X + Y$ y compruebe su respuesta.

2.75. Sean X y Y variables aleatorias independientes distribuidas de manera idéntica con función de densidad
$$f(t) = \begin{cases} e^{-t} & t \ge 0 \\ 0 & \text{si no es así} \end{cases}$$
Encuentre la función de densidad de $X + Y$ y compruebe su respuesta.

2.76. Repita el problema 2.21 haciendo primero la transformación $2Y = Z$ y use después convoluciones para hallar la función de densidad de $U = X + Z$.

2.77. Si las variables aleatorias independientes X_1 y X_2 están distribuidas de manera idéntica con función de densidad
$$f(t) = \begin{cases} te^{-t} & t \ge 0 \\ 0 & t < 0 \end{cases}$$
encuentre la función de densidad de $X_1 + X_2$.

APLICACIONES DE LA PROBABILIDAD GEOMÉTRICA

2.78. En un segmento de recta cuya longitud es $a > 0$ se van a elegir dos puntos al azar. Encuentre la probabilidad de que los tres segmentos de recta que se forman sean los lados de un triángulo.

2.79. Se sabe que un autobús llega a cierto lugar en algún momento al azar, entre las 3:00 p.m. y las 3:30 p.m. Una persona decide llegar a este lugar en algún momento al azar entre esas horas y esperar el autobús a lo más 5 minutos. Si pierde el autobús tomará el tren subterráneo. ¿Cuál es la probabilidad de que tome el tren subterráneo?

2.80. Las longitudes de dos segmentos de recta AB y CD son 8 y 6 unidades, respectivamente. Sobre AB y CD se van a elegir, respectivamente, dos puntos P y Q al azar. Muestre que la probabilidad de que: el área de un triángulo que tendrá como altura AP y como base CQ y será mayor de 12 unidades cuadradas, es igual a $(1 - \ln 2)/2$.

PROBLEMAS DIVERSOS

2.81. Suponga que $f(x) = c/3^x$, $x = 1, 2, \ldots$, es la función de probabilidad de una variable aleatoria X. *a*) Determine c. *b*) Encuentre la función de distribución. *c*) Represente mediante una gráfica la función de probabilidad y la función de distribución. *d*) Encuentre $P(2 \le X < 5)$. *e*) Encuentre $P(X \ge 3)$.

2.82. Suponga que
$$f(x) = \begin{cases} cxe^{-2x} & x \ge 0 \\ 0 & \text{si no es así} \end{cases}$$
es la función de densidad de una variable aleatoria X. *a*) Determine c. *b*) Encuentre la función de distribución. *c*) Represente en forma gráfica la función de densidad y la función de distribución. *d*) Encuentre $P(X \ge 1)$. *e*) Encuentre $P(2 \le X < 3)$.

2.83. La función de probabilidad de una variable aleatoria X está dada por

$$f(x) = \begin{cases} 2p & x = 1 \\ p & x = 2 \\ 4p & x = 3 \\ 0 & \text{si no es así} \end{cases}$$

donde p es una constante. Encuentre *a*) $P(0 \leq X < 3)$, *b*) $P(X > 1)$.

2.84. *a*) Demuestre que para una constante c adecuada

$$F(x) = \begin{cases} 0 & x \leq 0 \\ c(1 - e^{-x})^2 & x > 0 \end{cases}$$

es la función de distribución de una variable aleatoria X y encuentre esta c. *b*) Determine $P(1 < X < 2)$.

2.85. Una variable aleatoria X tiene la función de densidad

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre la función de densidad de la variable $Y = X^2$ y verifique su respuesta.

2.86. Dos variables aleatorias independientes, X y Y tienen, respectivamente, las funciones de densidad

$$f(x) = \begin{cases} c_1 e^{-2x} & x > 0 \\ 0 & x \leq 0 \end{cases} \qquad g(y) = \begin{cases} c_2 y e^{-3y} & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Encuentre *a*) c_1 y c_2, *b*) $P(X + Y > 1)$, *c*) $P(1 < X < 2, Y \geq 1)$, *d*) $P(1 < X < 2)$, *e*) $P(Y \geq 1)$.

2.87. En el problema 2.86, ¿cuál es la relación entre las respuestas de los incisos *c*), *d*) y *e*)? Justifique su respuesta.

2.88. Sean X y Y variables aleatorias cuya función de densidad conjunta es

$$f(x, y) = \begin{cases} c(2x + y) & 0 < x < 1, 0 < y < 2 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) la constante c, *b*) $P(X > \frac{1}{2}, Y < \frac{3}{2})$, *c*) la función de densidad (marginal) de X, *d*) la función de densidad (marginal) de Y.

2.89. En el problema 2.88, ¿es $P(X > \frac{1}{2}, Y < \frac{3}{2}) = P(X > \frac{1}{2})P(Y < \frac{3}{2})$? ¿Por qué?

2.90. En el problema 2.86 encuentre la función de densidad *a*) de X^2, *b*) de $X + Y$.

2.91. La función de densidad conjunta de X y Y es

$$f(x, y) = \begin{cases} 1/y & 0 < x < y, 0 < y < 1 \\ 0 & \text{si no es así} \end{cases}$$

a) Determine si X y Y son independientes. *b*) Encuentre $P(X > \frac{1}{2})$. *c*) Encuentre $P(X < \frac{1}{2}, Y > \frac{1}{3})$. *d*) Encuentre $P(X + Y > \frac{1}{2})$.

2.92. Generalice *a*) el problema 2.74 y *b*) el problema 2.75 a tres o más variables.

2.93. Sean X y Y variables aleatorias independientes distribuidas de manera idéntica y con función de densidad $f(u) = (2\pi)^{-1/2}e^{-u^2/2}$, $-\infty < u < \infty$. Encuentre la función de densidad de $Z = X^2 + Y^2$.

2.94. La función de probabilidad conjunta de las variables X y Y se da en la tabla 2-9. *a*) Encuentre las funciones de probabilidad marginal de X y Y. *b*) Encuentre $P(1 \leq X < 3, Y \geq 1)$. *c*) Determine si X y Y son independientes.

Tabla 2-9

X \ Y	0	1	2
0	1/18	1/9	1/6
1	1/9	1/18	1/9
2	1/6	1/6	1/18

2.95. Suponga que la función de probabilidad conjunta de las variables aleatorias X y Y está dada por

$$f(x, y) = \begin{cases} cxy & 0 \leq x \leq 2, 0 \leq y \leq x \\ 0 & \text{si no es así} \end{cases}$$

a) Determine si X y Y son independientes. *b*) Encuentre $P(\frac{1}{2} < X < 1)$. *c*) Encuentre $P(Y \geq 1)$. *d*) Encuentre $P(\frac{1}{2} < X < 1, Y \geq 1)$.

2.96. Sean X y Y variables aleatorias independientes cada una con función de densidad

$$f(u) = \frac{\lambda^u e^{-\lambda}}{u} \qquad u = 0, 1, 2, \ldots$$

donde $\lambda > 0$. Demuestre que la función de densidad de $X + Y$ es

$$g(u) = \frac{(2\lambda)^u e^{-2\lambda}}{u!} \qquad u = 0, 1, 2, \ldots$$

2.97. Se debe partir una vara de longitud L en dos partes. ¿Cuál es la probabilidad de que la longitud de una de las partes sea más del doble que la de la otra? Proporcione de manera clara los supuestos que deben hacerse. Analice si considera que estos supuestos son realistas y cómo pueden mejorarse si no lo son.

2.98. Un piso está formado por cuadrados de lados l. Una aguja de longitud $a < l$ se lanza sobre el piso. Demostrar que la probabilidad de que la aguja cruce por lo menos un lado es igual a $a(4l - a)/\pi l^2$.

2.99. Dada la longitud de una aguja, ¿cuál debe ser la longitud de los cuadrados del problema 2.98 para que la probabilidad de intersección sea máxima? Explique su respuesta.

2.100. Sea

$$f(x, y, z) = \begin{cases} 24xy^2z^3 & 0 < x < 1, 0 < y < 1, 0 < z < 1 \\ 0 & \text{si no es así} \end{cases}$$

variables aleatorias X, Y y Z. Encuentre *a*) $P(X > \frac{1}{2}, Y < \frac{1}{2}, Z > \frac{1}{2})$, *b*) $P(Z < X + Y)$.

2.101. Una corriente cilíndrica de partículas, de radio a, se dirige hacia un objetivo hemisférico ABC con centro en O como se muestra en la figura 2-26. Suponga que la distribución de las partículas está dada por

$$f(r) = \begin{cases} 1/a & 0 < r < a \\ 0 & \text{si no es así} \end{cases}$$

donde r es la distancia desde el eje OB. Demuestre que la distribución de las partículas sobre el objetivo está dada por

$$g(\theta) = \begin{cases} \cos\theta & 0 < \theta < \pi/2 \\ 0 & \text{si no es así} \end{cases}$$

donde θ es el ángulo que forma la recta OP (de O a un punto cualquiera P sobre el objetivo) con el eje.

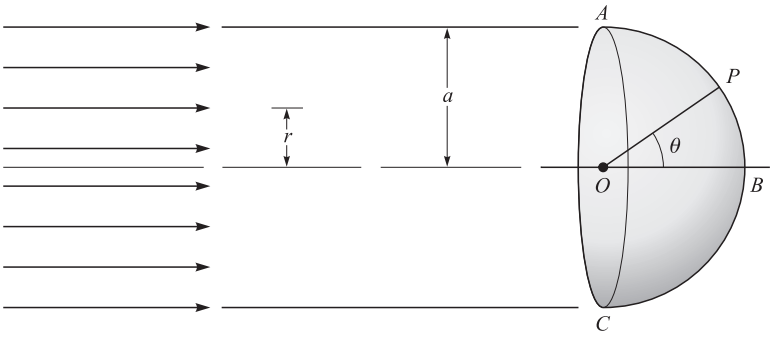


Figura 2-26

2.102. En el problema 2.101, encuentre la probabilidad de que una partícula choque con el objetivo entre $\theta = 0$ y $\theta = \pi/4$.

2.103. Suponga que las variables aleatorias X, Y y Z tienen función de densidad conjunta

$$f(x, y, z) = \begin{cases} 1 - \cos\pi x \cos\pi y \cos\pi z & 0 < x < 1, 0 < y < 1, 0 < z < 1 \\ 0 & \text{si no es así} \end{cases}$$

Demuestre que aunque cualesquiera dos de estas variables aleatorias son independientes, las tres no lo son.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

	/	/	/	/

x	0	1	2
$f(x)$	3/28	15/28	5/14

x	0	1	2
$f(x)$	9/64	15/32	25/64

x	0	1	2	3	4
$f(x)$	$\frac{194\,580}{270\,725}$	$\frac{69\,184}{270\,725}$	$\frac{6\,768}{270\,725}$	$\frac{192}{270\,725}$	$\frac{1}{270\,725}$

x	0	1	2	3
$f(x)$	1/8	1/2	7/8	1

a)

x	1	2	3	4
$f(x)$	1/8	1/4	3/8	1/4

b) 3/4 *c*) 7/8 *d*) 3/8 *e*) 7/8

2.47. *a*) 3 *b*) $e^{-3} - e^{-6}$ *c*) e^{-9} *d*) $1 - e^{-3}$ **2.48.** $F(x) = \begin{cases} 1 - e^{-3x} & x \geq 0 \\ 0 & x \leq 0 \end{cases}$

2.49. *a*) 6/29 *b*) 15/29 / $F(x) = \begin{cases} 0 & x \leq 1 \\ (2x^3 - 2)/29 & 1 \leq x \leq 2 \\ (3x^2 + 2)/29 & 2 \leq x \leq 3 \\ 1 & x \geq 3 \end{cases}$

$f(x) = \begin{cases} x^2/9 & 0 \leq x < 3 \\ 0 & \text{si no es así} \end{cases}$ / /

/ / / $F(x) = \begin{cases} 0 & x \leq 0 \\ x^2/4 & 0 \leq x \leq 2 \\ 1 & x \geq 2 \end{cases}$

/ / / / / / /

$f_1(x) = \begin{cases} x/6 & x = 1, 2, 3 \\ 0 & \text{otro valor de } x \end{cases}$ $f_2(y) = \begin{cases} y/6 & y = 1, 2, 3 \\ 0 & \text{otro valor de } y \end{cases}$

/ / / /

$F_1(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{2}(x^3 + x) & 0 \leq x \leq 1 \\ 1 & x \geq 1 \end{cases}$ $F_2(y) = \begin{cases} 0 & y \leq 0 \\ \frac{1}{2}(y^3 + y) & 0 \leq y \leq 1 \\ 1 & y \geq 1 \end{cases}$

$f(x \mid y) = f_1(x)$ = 1, 2, 3 (vea el problema 2.55)

$f(y \mid x) = f_2(y)$ = 1, 2, 3 (vea el problema 2.55)

2.59. $f(x|y) = \begin{cases} (x + y)/(y + \frac{1}{2}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otro valor de } x, 0 \leq y \leq 1 \end{cases}$

$f(y \mid x) = \begin{cases} (x + y)/(x + \frac{1}{2}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & 0 \leq x \leq 1, \text{ otro valor de } y \end{cases}$

$f(x \mid y) = \begin{cases} (x^2 + y^2)/(y^2 + \frac{1}{3}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otro valor de } x\,, 0 \leq y \leq 1 \end{cases}$

$f(y \mid x) = \begin{cases} (x^2 + y^2)/(x^2 + \frac{1}{3}) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & 0 \leq x \leq 1, \text{ otro valor de } y \end{cases}$

$f(x \mid y) = \begin{cases} e^{-x} & x \geq 0, y \geq 0 \\ 0 & x < 0, y \geq 0 \end{cases}$ $f(y \mid x) = \begin{cases} e^{-y} & x \geq 0, y \geq 0 \\ 0 & x \geq 0, y < 0 \end{cases}$

$e^{-\sqrt{y}}/2\sqrt{y}$ para $y > 0$; 0 si no es así $(2\pi)^{-1/2} y^{-1/2} e^{-y/2}$ >

$1/\pi$ $-\pi/2 < y < \pi/2$; 0

$g(y) = \begin{cases} \frac{1}{6} & -5 < y < 1 \\ 0 & \text{si no es así} \end{cases}$ $g(y) = \begin{cases} \frac{1}{6}(1 - y)^{-2/3} & 0 < y < 1 \\ \frac{1}{6}(y - 1)^{-2/3} & 1 < y < 2 \\ 0 & \text{si no es así} \end{cases}$

$ve^{-v}/(1 + u)^2$ $\geq 0, v \geq 0$; 0

2.73. $g(z) = \begin{cases} -\ln z & 0 < z < 1 \\ 0 & \text{si no es así} \end{cases}$ **2.77.** $g(x) = \begin{cases} x^3 e^{-x}/6 & x \geq 0 \\ 0 & x < 0 \end{cases}$

2.74. $g(u) = \begin{cases} u & 0 \leq u \leq 1 \\ 2 - u & 1 \leq u \leq 2 \\ 0 & \text{si no es así} \end{cases}$ **2.78.** $1/4$

2.75. $g(u) = \begin{cases} ue^{-u} & u \geq 0 \\ 0 & u < 0 \end{cases}$ **2.79.** $61/72$

2.81. *a*) 2 *b*) $F(x) = \begin{cases} 0 & x < 1 \\ 1 - 3^{-y} & y \leq x < y + 1;\ y = 1, 2, 3, \ldots \end{cases}$ *d*) $26/81$ *e*) $1/9$

2.82. *a*) 4 *b*) $F(x) = \begin{cases} 1 - e^{-2x}(2x + 1) & x \geq 0 \\ 0 & x < 0 \end{cases}$ *d*) $3e^{-2}$ *e*) $5e^{-4} - 7e^{-6}$

2.83. *a*) $3/7$ *b*) $5/7$ **2.84.** *a*) $c = 1$ *b*) $e^{-4} - 3e^{-2} + 2e^{-1}$

2.86. *a*) $c_1 = 2, c_2 = 9$ *b*) $9e^{-2} - 14e^{-3}$ *c*) $4e^{-5} - 4e^{-7}$ *d*) $e^{-2} - e^{-4}$ *e*) $4e^{-3}$

2.88. *a*) $1/4$ *b*) $27/64$ *c*) $f_1(x) = \begin{cases} x + \frac{1}{2} & 0 < x < 1 \\ 0 & \text{si no es así} \end{cases}$ *d*) $f_2(y) = \begin{cases} \frac{1}{4}(y + 1) & 0 < y < 2 \\ 0 & \text{si no es así} \end{cases}$

2.90. *a*) $\begin{cases} e^{-2y}/\sqrt{y} & y > 0 \\ 0 & \text{si no es así} \end{cases}$ *b*) $\begin{cases} 18e^{-2u} & u > 0 \\ 0 & \text{si no es así} \end{cases}$

2.91. *b*) $\frac{1}{2}(1 - \ln 2)$ *c*) $\frac{1}{6} + \frac{1}{2}\ln 2$ *d*) $\frac{1}{2}\ln 2$ **2.95.** *b*) $15/256$ *c*) $9/16$ *d*) 0

2.93. $g(z) = \begin{cases} \frac{1}{2}e^{-z/2} & z \geq 0 \\ 0 & z < 0 \end{cases}$ **2.100.** *a*) $45/512$ *b*) $1/14$

2.94. *b*) $7/18$ **2.102.** $\sqrt{2}/2$

PROBLEMAS APORTADOS

1. Se desea formar un comité de dos estudiantes seleccionados al azar entre tres cursos diferentes. Sólo hay tres estudiantes elegibles en el curso de estadística, dos en el curso de economía y tres en física. Si X es el número de estudiantes seleccionados en el curso de estadística y Y es el número de estudiantes seleccionados en el curso de economía, construya una tabla que muestre los valores de la distribución de probabilidad conjunta de X y de Y.

2. En una refinería se hace un estudio de los niveles de contaminación que se presentan en el lugar donde está ubicada la refinería. Para ello se diseña un experimento para evaluar el contenido de SO_2 (dióxido de azufre) en las salidas de las chimeneas. La refinería cuenta con 18 chimeneas y la evaluación se hace en 9 seleccionadas en forma aleatoria. El superintendente de la refinería sospecha que 6 chimeneas contaminan. Al realizar la evaluación se desea comprobar que al menos 4 de las 6 chimeneas contaminan y que la contaminación sea mayor a 30%.

3. La función de densidad conjunta de dos variables aleatorias continuas x y y es

$$f(x,y) = \begin{cases} cx(1+3y^2), & 0 < x < 4, \quad 0 < y < 2 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

a) Encontrar el valor de la constante c.
b) Encontrar $P(1 < X < 2, 0 < Y < 1)$.
c) Encontrar la función de distribución marginal de X.
d) Encontrar la función de distribución marginal de Y.

CAPÍTULO 3

Esperanza matemática

DEFINICIÓN DE ESPERANZA MATEMÁTICA

Un concepto muy importante en probabilidad y estadística es el de la *esperanza matemática*, *valor esperado* o, simplemente, *esperanza*, de una variable aleatoria. Dada una variable aleatoria discreta X cuyos posibles valores son x_1, x_2, . . . , x_n, la esperanza de X se define como

$$E(X) = x_1P(X = x_1) + \cdots + x_nP(X = x_n) = \sum_{j=1}^{n} x_jP(X = x_j) \tag{1}$$

o, de manera, equivalente, si $P(X = x_j) = f(x_j)$,

$$E(X) = x_1f(x_1) + \cdots + x_nf(x_n) = \sum_{j=1}^{n} x_jf(x_j) = \sum xf(x) \tag{2}$$

donde la última suma se lleva a cabo con todos los valores apropiados de x. Como un caso especial de (2), en el que todas las probabilidades son iguales, se tiene

$$E(X) = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{3}$$

que se conoce como la *media aritmética* o simplemente la *media* de $x_1, x_2, \ldots, x_n$.

Si X toma un número infinito de valores $x_1, x_2, \ldots$, entonces $E(X) = \sum_{j=1}^{\infty} x_j f(x_j)$ siempre que esta sucesión infinita converja de manera absoluta.

Dada una variable aleatoria continua X con una función de densidad $f(x)$, la esperanza de X se define como

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx \tag{4}$$

siempre que esta integral converja de manera absoluta.

A la esperanza de X suele llamársele *media* de X y se le denota como μ_X o, simplemente μ, cuando se sobreentiende la variable aleatoria de que se trata.

La media o esperanza de X da un solo valor que funciona como representante o promedio de los valores de X, razón por la cual suele considerársele como una *medida de tendencia central*. En la página 83 se consideran otras medidas.

EJEMPLO 3.1 Suponga que se va a efectuar un juego con un solo dado que se considera no está cargado. En este juego, el jugador gana \$20 si obtiene 2, \$40 si obtiene 4 y pierde \$30 si éste es un 6; si obtiene cualquier otro número ni gana ni pierde. Encuentre la suma esperada de dinero que puede ganar.

Sea X la variable aleatoria que da la cantidad de dinero ganada en un lanzamiento. Las cantidades que se pueden ganar cuando cae un 1, 2,…, 6 son $x_1, x_2, \ldots, x_6$, respectivamente, y las probabilidades de éstas son $f(x_1), f(x_2), \ldots, f(x_6)$. En la tabla 3-1 se presenta la función de probabilidad de X. Por tanto, el valor esperado o esperanza es

$$E(X) = (0)\left(\frac{1}{6}\right) + (20)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (40)\left(\frac{1}{6}\right) + (0)\left(\frac{1}{6}\right) + (-30)\left(\frac{1}{6}\right) = 5$$

Tabla 3-1

x_j	0	+20	0	+40	0	−30
$f(x_j)$	1/6	1/6	1/6	1/6	1/6	1/6

Se concluye que el jugador puede esperar ganar \$5. Por tanto, si es un juego legal, el jugador puede esperar pagar \$5 por jugar.

EJEMPLO 3.2 La función de densidad de una variable aleatoria X está dada por

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & \text{si no es así} \end{cases}$$

El valor esperado de X es, entonces

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_0^2 x\left(\frac{1}{2}x\right)dx = \int_0^2 \frac{x^2}{2}\,dx = \frac{x^3}{6}\Big|_0^2 = \frac{4}{3}$$

FUNCIONES DE VARIABLES ALEATORIAS

Sea X una variable aleatoria discreta cuya función de probabilidad es $f(x)$. Entonces, $Y = g(X)$ también es una variable aleatoria discreta, y la función de probabilidad de Y es

$$h(y) = P(Y = y) = \sum_{\{x|g(x)=y\}} P(X = x) = \sum_{\{x|g(x)=y\}} f(x)$$

Si X toma los valores $x_1, x_2, \ldots, x_n$ y Y los valores $y_1, y_2, \ldots, y_m$ $(m \le n)$, entonces $y_1h(y_1) + y_2h(y_2) + \cdots + y_mh(y_m) = g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n)$. Por tanto,

$$\begin{aligned} E[g(X)] &= g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n) \\ &= \sum_{j=1}^{n} g(x_j)f(x_j) = \sum g(x)f(x) \end{aligned} \tag{5}$$

De manera similar, sea X una variable aleatoria continua cuya densidad de probabilidad es $f(x)$, entonces puede demostrarse que

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx \tag{6}$$

Observe que ni en (5) ni en (6) intervienen, respectivamente, ni la función de probabilidad ni la función de densidad de probabilidad de $Y = g(X)$.

Es fácil hacer generalizaciones a funciones de dos o más variables aleatorias. Por ejemplo, si X y Y son dos variables aleatorias continuas cuya función de densidad conjunta es $f(x, y)$, la esperanza de $g(X, Y)$ está dada por

$$E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f(x, y)\,dx\,dy \tag{7}$$

EJEMPLO 3.3 Si X es la variable aleatoria del ejemplo 3.2,

$$E(3X^2 - 2X) = \int_{-\infty}^{\infty} (3x^2 - 2x)f(x)\,dx = \int_0^2 (3x^2 - 2x)\left(\frac{1}{2}x\right)dx = \frac{10}{3}$$

ALGUNOS TEOREMAS SOBRE LA ESPERANZA

Teorema 3-1 Si c es una constante cualquiera, entonces

$$E(cX) = cE(X) \tag{8}$$

Teorema 3-2 Si X y Y son variables aleatorias de cualquier tipo, entonces

$$E(X + Y) = E(X) + E(Y) \tag{9}$$

Teorema 3-3 Si X y Y son variables aleatorias independientes, entonces

$$E(XY) = E(X)E(Y) \tag{10}$$

Las generalizaciones de estos teoremas son sencillas.

LA VARIANZA Y LA DESVIACIÓN ESTÁNDAR

En la página 75 se hizo notar que a la esperanza de una variable aleatoria suele llamársele la *media* y denotarla como μ. Otra cantidad muy importante en probabilidad y estadística es la *varianza* que se define así

$$\text{Var}(X) = E[(X - \mu)^2] \tag{11}$$

La varianza es un número no negativo. A la raíz cuadrada positiva de la varianza se le llama la *desviación estándar*, y está dada por

$$\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{E[(X - \mu)^2]} \tag{12}$$

En los casos en que no pueda haber confusión, la desviación estándar se denota como σ en lugar de σ_X, y la varianza como σ^2.

Sea X una variable aleatoria discreta que tome los valores $x_1, x_2, \ldots, x_n$ cuya función de probabilidad sea $f(x)$, entonces la varianza está dada por

$$\sigma_X^2 = E[(X - \mu)^2] = \sum_{j=1}^{n}(x_j - \mu)^2 f(x_j) = \sum(x - \mu)^2 f(x) \tag{13}$$

En el caso especial en el que todas las probabilidades sean iguales, (13) se convierte en

$$\sigma^2 = [(x_1 - \mu)^2 + (x_2 - \mu)^2 + \cdots + (x_n - \mu)^2]/n \tag{14}$$

que es la varianza de un conjunto de n números $x_1, x_2, \ldots, x_n$.

Si X toma un número infinito de valores $x_1, x_2, \ldots$, entonces $\sigma_X^2 = \sum_{j=1}^{\infty}(x_j - \mu)^2 f(x_j)$, siempre que esta sucesión converja.

Si X es una variable aleatoria continua cuya función de densidad es $f(x)$, entonces la varianza está dada por

$$\sigma_X^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)\,dx \tag{15}$$

siempre que la integral converja.

La varianza (o la desviación estándar) es una medida de la *dispersión* o *variación* de los valores de la variable aleatoria alrededor de la media μ. Si los valores tienden a concentrarse cerca de la media, la varianza es pequeña, mientras que si tienden a distribuirse alejados de la media, la varianza es grande. Estas situaciones se representan de manera gráfica en la figura 3-1 en el caso de dos distribuciones continuas que tienen la misma media μ.

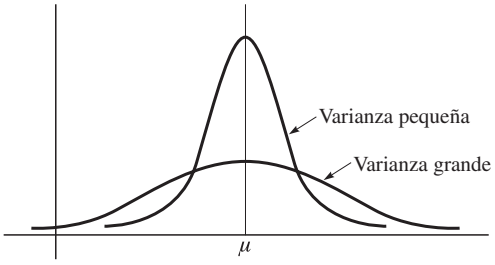


Figura 3-1

EJEMPLO 3.4 Encuentre la varianza y la desviación estándar de la variable aleatoria del ejemplo 3.2. En él se determinó que la media es $\mu = E(X) = 4/3$. En consecuencia, la varianza es

$$\sigma^2 = E\left[\left(X - \frac{4}{3}\right)^2\right] = \int_{-\infty}^{\infty}\left(x - \frac{4}{3}\right)^2 f(x)\,dx = \int_0^2\left(x - \frac{4}{3}\right)^2\left(\frac{1}{2}x\right)dx = \frac{2}{9}$$

y, por tanto, la desviación estándar es $\sigma = \sqrt{\frac{2}{9}} = \frac{\sqrt{2}}{3}$.

Observe que si X tiene alguna *dimensión* o *unidades*, como por ejemplo, *centímetros* (cm), entonces la varianza de X tiene como unidades cm^2, mientras que la desviación estándar tiene las mismas unidades que X, o sea, cm. A esto se debe que la desviación estándar se usa con mayor frecuencia.

ALGUNOS TEOREMAS SOBRE LA VARIANZA

Teorema 3-4 $$\sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2 \tag{16}$$

donde $\mu = E(X)$.

Teorema 3-5 Si c es una constante cualquiera,

$$\text{Var}(cX) = c^2\,\text{Var}(X) \tag{17}$$

Teorema 3-6 La cantidad $E[(X - a)^2]$ es mínima cuando $a = \mu^2 = E(X)$.

Teorema 3-7 Si X y Y son variables aleatorias independientes,

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \quad \text{o bien} \quad \sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 \tag{18}$$

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) \quad \text{o bien} \quad \sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 \tag{19}$$

El teorema 3-7 puede generalizarse con facilidad a más de dos variables independientes. En palabras, la varianza de una suma de variables independientes es igual a la suma de sus varianzas.

VARIABLES ALEATORIAS ESTANDARIZADAS

Dada una variable aleatoria X con media μ y desviación estándar σ ($\sigma > 0$), puede definirse una correspondiente *variable aleatoria estandarizada* dada por

$$X^* = \frac{X - \mu}{\sigma} \tag{20}$$

Una propiedad importante de X^* es que tiene media cero y varianza 1, lo que explica que se le llame *estandarizada*, es decir,

$$E(X^*) = 0, \qquad \text{Var}(X^*) = 1 \tag{21}$$

A los valores de una variable estandarizada suele llamárseles *puntuaciones estándar* y X se expresa en *unidades estándar* (es decir, σ se toma como la unidad para medir $X - \mu$).

Las variables estandarizadas sirven para comparar distribuciones diferentes.

MOMENTOS

El *r-ésimo momento de una variable aleatoria X alrededor de la media* μ, se le nombra *r-ésimo momento central*, se define como

$$\mu_r = E[(X - \mu)^r] \tag{22}$$

donde $r = 0, 1, 2, \ldots$. De ello se deduce que $\mu_0 = 1$, $\mu_1 = 0$ y $\mu_2 = \sigma^2$, es decir, el segundo momento central o segundo momento alrededor de la media es la varianza. Se tiene, suponiendo que la convergencia es absoluta,

$$\mu_r = \sum (x - \mu)^r f(x) \qquad \text{(variable discreta)} \tag{23}$$

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x)\, dx \qquad \text{(variable continua)} \tag{24}$$

El *r-ésimo momento de X alrededor del origen*, también se le conoce como *r-ésimo momento en bruto*, se define como

$$\mu_r' = E(X^r) \tag{25}$$

donde $r = 0, 1, 2, \ldots$, para este caso existen fórmulas análogas a (23) y (24) en las que $\mu = 0$.

La relación entre estos momentos está dada por

$$\mu_r = \mu_r' - \binom{r}{1}\mu_{r-1}'\mu + \cdots + (-1)^j\binom{r}{j}\mu_{r-j}'\mu^j + \cdots + (-1)^r\mu_0'\mu^r \tag{26}$$

Como caso especial se tiene, con $\mu_1' = \mu$ y $\mu_0' = 1$.

$$\begin{aligned} \mu_2 &= \mu_2' - \mu^2 \\ \mu_3 &= \mu_3' - 3\mu_2'\mu + 2\mu^3 \\ \mu_4 &= \mu_4' - 4\mu_3'\mu + 6\mu_2'\mu^2 - 3\mu^4 \end{aligned} \tag{27}$$

FUNCIONES GENERADORAS DE MOMENTOS

La *función generadora de momentos* de X se define como

$$M_X(t) = E(e^{tX}) \tag{28}$$

es decir, si se supone convergencia,

$$M_X(t) = \sum e^{tx} f(x) \qquad \text{(variable discreta)} \tag{29}$$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \qquad \text{(variable continua)} \tag{30}$$

Puede demostrarse que la expansión de las series de Taylor es [problema 3.15*a*)]

$$M_X(t) = 1 + \mu t + \mu_2'\frac{t^2}{2!} + \cdots + \mu_r'\frac{t^r}{r!} + \cdots \tag{31}$$

La razón del nombre *función generadora de momentos* es que los coeficientes de esta expansión permiten hallar los momentos. A partir de esta expansión se puede demostrar que [problema 3.15*b*)]

$$\mu_r' = \frac{d^r}{dt^r} M_X(t)\bigg|_{t=0} \tag{32}$$

es decir, μ_r' es la r-ésima derivada de $M_X(t)$ evaluada en $t = 0$. Cuando no hay lugar a confusión, suele escribirse $M(t)$ en lugar de $M_X(t)$.

ALGUNOS TEOREMAS SOBRE FUNCIONES GENERADORAS DE MOMENTOS

Teorema 3-8 Si $M_X(t)$ es la función generadora de momentos de la variable aleatoria X y a y b ($b \neq 0$) son constantes, la función generadora de momentos de $(X + a)/b$ es

$$M_{(X+a)/b}(t) = e^{at/b} M_X\left(\frac{t}{b}\right) \tag{33}$$

Teorema 3-9 Si X y Y son variables aleatorias independientes cuyas funciones generadoras de momentos son $M_X(t)$ y $M_Y(t)$, respectivamente, entonces

$$M_{X+Y}(t) = M_X(t)\, M_Y(t) \tag{34}$$

El teorema 3-9 puede generalizarse con facilidad a más de dos variables aleatorias independientes. En palabras, la función generadora de momentos de una suma de variables aleatorias independientes es igual al producto de sus funciones generadoras de momentos.

Teorema 3-10 (**Teorema de unicidad**) Suponga que X y Y sean variables aleatorias cuyas funciones generadoras de momentos sean $M_X(t)$ y $M_Y(t)$, respectivamente. Entonces, X y Y tienen la misma distribución de probabilidad si y sólo si $M_X(t) = M_Y(t)$ de manera idéntica.

FUNCIONES CARACTERÍSTICAS

Si se hace $t = i\omega$, donde i es la unidad imaginaria, en la función generadora de momentos se obtiene una función importante conocida como *función característica.* Esta función se denota como

$$\phi_X(\omega) = M_X(i\omega) = E(e^{i\omega X}) \tag{35}$$

Se concluye que

$$\phi_X(\omega) = \sum e^{i\omega x} f(x) \qquad \text{(variable discreta)} \tag{36}$$

$$\phi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\, dx \qquad \text{(variable continua)} \tag{37}$$

Dado que $\left|e^{i\omega x}\right| = 1$, tanto la serie como la integral convergen siempre de manera absoluta.

Los resultados correspondientes a (31) y (32) son

$$\phi_X(\omega) = 1 + i\mu\omega - \mu_2' \frac{\omega^2}{2!} + \cdots + i^r \mu_r' \frac{\omega^r}{r!} + \cdots \tag{38}$$

donde

$$\mu_r' = (-1)^r i^r \frac{d^r}{d\omega^r} \phi_X(\omega)\bigg|_{\omega=0} \tag{39}$$

Cuando no hay lugar a confusión, suele escribirse $\phi(\omega)$ en lugar de $\phi_X(\omega)$.

Los teoremas de las funciones características correspondientes a los teoremas 3-8, 3-9 y 3-10 son los siguientes.

Teorema 3-11 Si $\phi_X(\omega)$ es la función característica de la variable aleatoria X y a y b ($b \neq 0$) son constantes, entonces la función característica de $(X + a)/b$ es

$$\phi_{(X+a)/b}(\omega) = e^{ai\omega/b} \phi_X\left(\frac{\omega}{b}\right) \tag{40}$$

Teorema 3-12 Si X y Y son variables aleatorias independientes cuyas funciones características son $\phi_X(\omega)$ y $\phi_Y(\omega)$, respectivamente, entonces

$$\phi_{X+Y}(\omega) = \phi_X(\omega)\, \phi_Y(\omega) \tag{41}$$

Es decir, la función característica de una suma de variables aleatorias independientes es igual al producto de sus funciones características.

Teorema 3-13 (**Teorema de unicidad**) Suponga que X y Y sean variables aleatorias cuyas funciones características sean $\phi_X(\omega)$ y $\phi_Y(\omega)$, respectivamente. Entonces X y Y tienen la misma distribución de probabilidad si y sólo si $\phi_X(\omega) = \phi_Y(\omega)$ de manera idéntica.

Una razón importante para introducir la función característica es que (37) representa la *transformada de Fourier* de la función de densidad $f(x)$. De acuerdo con la teoría de las transformadas de Fourier, la función de densidad puede determinarse con facilidad a partir de la función característica. En efecto,

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega x}\phi_X(\omega)\,d\omega \tag{42}$$

lo que suele conocerse como la *fórmula de inversión* o *transformada inversa de Fourier.* De manera similar, en el caso discreto puede demostrarse que la función de probabilidad $f(x)$ puede obtenerse de (36) usando las *series de Fourier*, que en el caso discreto es la análoga a la integral de Fourier. Vea el problema 3.39. Otra razón para usar la función característica es que ésta siempre existe, mientras que la función generadora de momentos quizá no exista.

VARIANZA DE DISTRIBUCIONES CONJUNTAS. COVARIANZA

Los resultados dados arriba para una variable pueden extenderse a dos o más variables. Por ejemplo, si X y Y son dos variables aleatorias continuas cuya función de densidad conjunta es $f(x, y)$, sus medias o esperanzas son

$$\mu_X = E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)\,dx\,dy, \qquad \mu_Y = E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)\,dx\,dy \tag{43}$$

y las varianzas son

$$\begin{aligned}\sigma_X^2 &= E[(X - \mu_X)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - \mu_X)^2 f(x, y)\,dx\,dy \\ \sigma_Y^2 &= E[(Y - \mu_Y)^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y - \mu_Y)^2 f(x, y)\,dx\,dy\end{aligned} \tag{44}$$

Observe que las funciones de densidad marginal de X y de Y no intervienen directamente ni en (43) ni en (44).

Otra cantidad que surge en el caso de dos variables X y Y es la *covarianza*, que se define como

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \tag{45}$$

En términos de la función de densidad conjunta $f(x, y)$, se tiene que

$$\sigma_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - \mu_X)(y - \mu_Y)f(x, y)\,dx\,dy \tag{46}$$

Observaciones similares pueden hacerse con respecto a dos variables aleatorias discretas. En tales casos, (43) y (46) se sustituyen por

$$\mu_X = \sum_x\sum_y xf(x, y) \qquad \mu_Y = \sum_x\sum_y yf(x, y) \tag{47}$$

$$\sigma_{XY} = \sum_x\sum_y (x - \mu_X)(y - \mu_Y)f(x, y) \tag{48}$$

donde las sumas consideran todos los valores discretos de X y Y.

Los siguientes son algunos teoremas importantes sobre covarianza.

Teorema 3-14 $$\sigma_{XY} = E(XY) - E(X)E(Y) = E(XY) - \mu_X\mu_Y \tag{49}$$

Teorema 3-15 Si X y Y son variables aleatorias independientes, entonces

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = 0 \tag{50}$$

Teorema 3-16 $$\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y) \tag{51}$$

o bien $$\sigma_{X\pm Y}^2 = \sigma_X^2 + \sigma_Y^2 \pm 2\sigma_{XY} \tag{52}$$

Teorema 3-17 $$|\sigma_{XY}| \le \sigma_X\sigma_Y \tag{53}$$

El inverso del teorema 3-15 no es necesariamente válido. Si X y Y son independientes, el teorema 3-16 se reduce al teorema 3-7.

COEFICIENTE DE CORRELACIÓN

Si X y Y son independientes, entonces $\text{Cov}(X, Y) = \sigma_{XY} = 0$. En otro caso, si X y Y son completamente dependientes, por ejemplo, cuando $X = Y$, entonces $\text{Cov}(X, Y) = \sigma_{XY} = \sigma_X\sigma_Y$. Esto conduce a una *medida de la dependencia* de las variables X y Y. Dada por

$$\rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \tag{54}$$

A ρ se le conoce como *coeficiente de correlación.* De acuerdo con el teorema 3-17 $-1 \leq \rho \leq 1$. Cuando $\rho = 0$ (es decir, en el que la covarianza sea cero), las variables X y Y *no están relacionadas*. Sin embargo, en tales casos las variables pueden o no ser independientes. En el capítulo 8 se analizan más casos de correlación.

ESPERANZA, VARIANZA Y MOMENTOS CONDICIONALES

Si X y Y tienen función de densidad conjunta $f(x, y)$, entonces, como se vio en el capítulo 2, la función de densidad condicional de Y dado X es $f(y \mid x) = f(x, y)/f_1(x)$ donde $f_1(x)$ es la función de densidad marginal de X. La *esperanza condicional* o *media condicional* de Y dado X se define como

$$E(Y \mid X = x) = \int_{-\infty}^{\infty} yf(y \mid x)\,dy \tag{55}$$

donde "$X = x$" se interpreta como $x < X \leq x + dx$ en el caso continuo. Los teoremas 3-1 y 3-2 también son válidos en el caso de la esperanza condicional.

Se apuntan las propiedades siguientes:

1. $E(Y \mid X = x) = E(Y)$ donde X y Y son independientes.
2. $E(Y) = \displaystyle\int_{-\infty}^{\infty} E(Y \mid X = x)f_1(x)\,dx.$

Suele ser útil calcular las esperanzas empleando la propiedad 2 en lugar de hacerlo directamente.

EJEMPLO 3.5 El tiempo promedio de viaje a una ciudad distante es c horas en automóvil y b horas en autobús. Una persona no sabe si ir en su automóvil o ir en autobús, por lo que decide lanzar una moneda. ¿Cuál es su tiempo esperado de viaje?

Aquí se tiene la distribución conjunta del resultado de lanzar una moneda, X, y del tiempo de viaje, Y, donde $Y = Y_{\text{automóvil}}$ si $X = 0$ y $Y = Y_{\text{autobús}}$ si $X = 1$. Se supone que $Y_{\text{automóvil}}$ y $Y_{\text{autobús}}$ son independientes de X, de manera que, de acuerdo con la propiedad 1 dada

$$E(Y \mid X = 0) = E(Y_{\text{automóvil}} \mid X = 0) = E(Y_{\text{automóvil}}) = c$$

y

$$E(Y \mid X = 1) = E(Y_{\text{autobús}} \mid X = 1) = E(Y_{\text{autobús}}) = b$$

En consecuencia, la propiedad 2 (sustituyendo la integral por una suma) da, en el caso en que se tenga una moneda no cargada,

$$E(Y) = E(Y \mid X = 0)P(X = 0) + E(Y \mid X = 1)P(X = 1) = \frac{c + b}{2}$$

De manera similar puede definirse la *varianza condicional* de Y dada X como

$$E[(Y - \mu_2)^2 \mid X = x] = \int_{-\infty}^{\infty} (y - \mu_2)^2 f(y \mid x)\,dy \tag{56}$$

donde $\mu_2 = E(Y \mid X = x)$. También puede definirse el r-ésimo *momento condicional* de Y, alrededor a un valor a, dado X como

$$E[(Y - a)^r \mid X = x] = \int_{-\infty}^{\infty} (y - a)^r f(y \mid x)\,dy \tag{57}$$

Los teoremas usuales de la varianza y los momentos se extienden a la varianza y a los momentos condicionales.

DESIGUALDAD DE CHEBYSHEV

Un teorema importante en probabilidad y estadística que revela una propiedad general de las variables aleatorias (discretas o continuas) que tienen media y varianza finitas es el que se conoce como *desigualdad de Chebyshev*.

Teorema 3-18 (**Desigualdad de Chebyshev**) Suponga que X es una variable aleatoria (discreta o continua) con media μ y varianza σ^2, ambas finitas. Entonces, si ϵ es cualquier número positivo,

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \tag{58}$$

o, con $\epsilon = k\sigma$,

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \tag{59}$$

EJEMPLO 3.6 Haciendo $k = 2$ en la desigualdad de Chebyshev (59), se ve que

$$P(|X - \mu| \geq 2\sigma) \leq 0.25 \qquad \text{o bien} \qquad P(|X - \mu| < 2\sigma) \geq 0.75$$

En palabras, la probabilidad de que X difiera de su media en más de dos desviaciones estándar es menor o igual a 0.25; de manera equivalente, la probabilidad de que X se encuentre a no más de dos desviaciones estándar de su media es mayor o igual a 0.75. Esto resulta muy interesante en vista del hecho de que ni siquiera se ha especificado la distribución de probabilidad de X.

LEY DE LOS GRANDES NÚMEROS

El teorema siguiente, llamado *ley de los grandes números*, es una interesante consecuencia de la desigualdad de Chebyshev.

Teorema 3-19 (**Ley de los grandes números**) Sean $X_1, X_2, \ldots, X_n$, variables aleatorias mutuamente independientes (discretas o continuas), cada una con media μ y varianza σ^2 finitas. Entonces, si $S_n = X_1 + X_2 + \cdots + X_n (n = 1, 2, \ldots)$,

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0 \tag{60}$$

Dado que S_n/n es la media aritmética de $X_1, X_2, \ldots, X_n$, este teorema afirma que la probabilidad de que la media aritmética S_n/n difiera de su valor esperado μ, en más de ϵ, tiende a cero a medida que $n \to \infty$. Un resultado más fuerte, que podría esperarse que fuera válido, es que $\lim_{n\to\infty} S_n/n = \mu$, pero esto en realidad es falso. Sin embargo, puede demostrarse que $\lim_{n\to\infty} S_n/n = \mu$ *con probabilidad uno.* Este resultado suele conocerse como *ley fuerte de los grandes números*, mientras que al teorema 3-19 se le conoce como *ley débil de los grandes números.* Cuando se habla de la "ley de los grandes números" sin calificativo, se hace referencia a la ley débil.

OTRAS MEDIDAS DE TENDENCIA CENTRAL

Como se vio, la media, o esperanza, de una variable aleatoria X proporciona una medida de la tendencia central de los valores de una distribución. Aunque la media es la más usada, también se emplean otras dos medidas de tendencia central. Éstas son la *moda* y la *mediana*.

1. **MODA.** La *moda* de una variable aleatoria discreta es aquel valor que se presenta con más frecuencia o, en otras palabras, tiene la mayor probabilidad de ocurrir. Algunas veces existen 2, 3 o más valores que tienen probabilidades relativamente grandes de ocurrencia. En tales casos, la distribución es *bimodal*, *trimodal* o *multimodal*, respectivamente. La moda de una variable aleatoria continua X es el valor (o los valores) de X en los que la función de densidad de probabilidad tiene un máximo relativo.

2. **MEDIANA.** La *mediana* es el valor x para el que $P(X < x) \leq \frac{1}{2}$ y $P(X > x) \leq \frac{1}{2}$. En el caso de una distribución continua se tiene $P(X < x) = \frac{1}{2} = P(X > x)$, y la mediana divide la curva de densidad en dos partes cuyas áreas son igual a $1/2$ cada una. En el caso de una distribución discreta puede no haber una mediana única (vea el problema 3.34).

PERCENTILES

Es útil subdividir el área bajo la curva de densidad mediante el uso de ordenadas de manera que el área a la izquierda de una ordenada sea algún porcentaje del área unitaria total. A los valores correspondientes a tales áreas se les llama *valores percentiles*, o simplemente *percentiles.* Así, por ejemplo, en la figura 3-2 el área a la izquierda de la ordenada x_α es α. Por ejemplo, el área a la izquierda de $x_{0.10}$ será 0.10, o bien 10%, y se le llamará el décimo percentil (conocido también como *primer decil*). La mediana es el *quincuagésimo percentil* (o *quinto decil*).

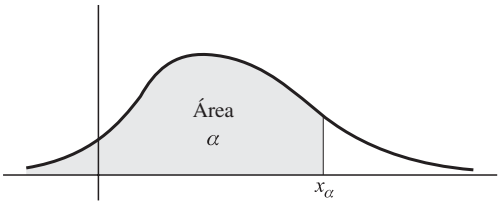


Figura 3-2

OTRAS MEDIDAS DE DISPERSIÓN

Así, como además de la media existen otras medidas de tendencia central, también además de la varianza y de la desviación estándar existen otras medidas de dispersión o variación de una variable aleatoria. Algunas de las más comunes son las siguientes.

1. **RANGO SEMIINTERCUARTIL.** Si $x_{0.25}$ y $x_{0.75}$ representan los valores del vigésimo quinto y del septuagésimo quinto percentiles, la diferencia $x_{0.75} - x_{0.25}$ se conoce como *rango intercuartil* y $\frac{1}{2}(x_{0.75} - x_{0.25})$ es el *rango semiintercuartil.*

2. **DESVIACIÓN MEDIA.** La *desviación media* (D.M.) de una variable aleatoria X se define como la esperanza de $|X - \mu|$, es decir, suponiendo que haya convergencia,

$$\text{D.M.}(X) = E[|X - \mu|] = \sum |x - \mu| f(x) \qquad \text{(variable discreta)} \tag{61}$$

$$\text{D.M.}(X) = E[|X - \mu|] = \int_{-\infty}^{\infty} |x - \mu| f(x)\, dx \qquad \text{(variable continua)} \tag{62}$$

SESGO Y CURTOSIS

1. **SESGO.** Con frecuencia, las distribuciones no son simétricas respecto de algún valor, sino que tienen una de sus colas más larga que la otra. Si la cola más larga se encuentra a la derecha, como en la figura 3-3, la distribución es *sesgada a la derecha*, y si la cola más larga se encuentra a la izquierda, como en la figura 3-4, es *sesgada a la izquierda.* A las medidas que describen esta asimetría se les llama *coeficientes de sesgo*, o simplemente *sesgo.* Una de estas medidas es

$$\alpha_3 = \frac{E[(X - \mu)^3]}{\sigma^3} = \frac{\mu_3}{\sigma^3} \tag{63}$$

La medida α_3 será positiva o negativa si la distribución es sesgada a la derecha o a la izquierda, respectivamente. En una distribución simétrica, $\alpha_3 = 0$.

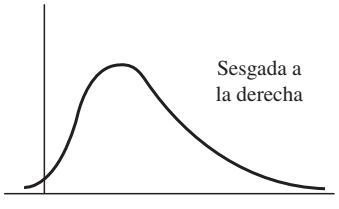


Figura 3-3

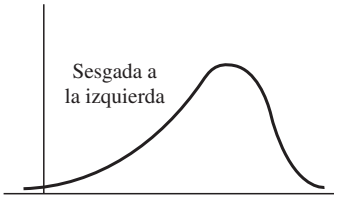


Figura 3-4

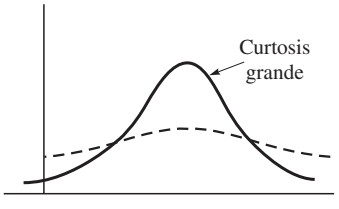


Figura 3-5

$$= \frac{E[(X-\mu)^4]}{\sigma^4} = \frac{\mu_4}{\sigma^4} \tag{64}$$

Esta medida acostumbra compararse con la de la curva normal (vea capítulo 4), que tiene un coeficiente de curtosis igual a 3. Vea también el problema 3.41.

PROBLEMAS RESUELTOS

ESPERANZA DE VARIABLES ALEATORIAS

3.1. Una lotería ofrece 200 premios de \$5, 20 premios de \$25 y 5 premios de \$100. Suponiendo que se van vender 10 000 boletos, ¿cuál es el precio justo que se debe pagar por uno de ellos?

Sea X una variable aleatoria que denote la cantidad de dinero que puede ganarse con un boleto. En la tabla 3-2 se muestran los valores de X junto con sus probabilidades. Por ejemplo, la probabilidad de obtener uno de los 20 boletos con los que se ganan \$25 es $20/10\,000 = 0.002$. Por tanto, la esperanza de X en dólares es

$$E(X) = (5)(0.02) + (25)(0.002) + (100)(0.0005) + (0)(0.9775) = 0.2$$

es decir 20 centavos. Por consiguiente, el precio justo a pagar por un boleto es 20 centavos. Sin embargo, como la lotería suele tener como objeto recaudar dinero, el precio del boleto será mayor.

Tabla 3-2

x (dólares)	5	25	100	0
$P(X = x)$	0.02	0.002	0.0005	0.9775

3.2. Encontrar la esperanza de la suma de los puntos que se obtienen al lanzar un par de dados no cargados.

Sean X y Y los puntos que aparecen al caer los dados. Se tiene

$$E(X) = E(Y) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + \cdots + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

En consecuencia, de acuerdo con el teorema 3-2,

$$E(X+Y) = E(X) + E(Y) = 7$$

3.3. Encontrar la esperanza de una variable aleatoria discreta X cuya función de probabilidad está dada por

$$f(x) = \left(\frac{1}{2}\right)^x \qquad (x = 1, 2, 3, \ldots)$$

Tenemos

$$E(X) = \sum_{x=1}^{\infty} x\left(\frac{1}{2}\right)^x = \frac{1}{2} + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{8}\right) + \cdots$$

Para hallar esta suma, sea $S = \frac{1}{2} + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{8}\right) + 4\left(\frac{1}{16}\right) + \cdots$

Entonces, $\frac{1}{2}S = \quad \frac{1}{4} + 2\left(\frac{1}{8}\right) + 3\left(\frac{1}{16}\right) + \cdots$

Cuando restamos, $\frac{1}{2}S = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 1$

Por tanto, $S = 2$.

3.4. Una variable aleatoria continua X tiene la densidad de probabilidad dada por

$$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Encontrar *a*) $E(X)$, *b*) $E(X^2)$.

a)
$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_{0}^{\infty} x(2e^{-2x})\,dx = 2\int_{0}^{\infty} xe^{-2x}\,dx$$
$$= 2\left[(x)\left(\frac{e^{-2x}}{-2}\right) - (1)\left(\frac{e^{-2x}}{4}\right)\right]\Bigg|_0^{\infty} = \frac{1}{2}$$

b)
$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = 2\int_{0}^{\infty} x^2 e^{-2x}\,dx$$
$$= 2\left[(x^2)\left(\frac{e^{-2x}}{-2}\right) - (2x)\left(\frac{e^{-2x}}{4}\right) + (2)\left(\frac{e^{-2x}}{-8}\right)\right]\Bigg|_0^{\infty} = \frac{1}{2}$$

3.5. La función de densidad conjunta de dos variables aleatorias X y Y está dada por

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{si no es así} \end{cases}$$

Encontrar *a*) $E(X)$, *b*) $E(Y)$, *c*) $E(XY)$, *d*) $E(2X + 3Y)$.

a)
$$E(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} x\left(\frac{xy}{96}\right)dx\,dy = \frac{8}{3}$$

b)
$$E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} y\left(\frac{xy}{96}\right)dx\,dy = \frac{31}{9}$$

c)
$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (xy)f(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} (xy)\left(\frac{xy}{96}\right)dx\,dy = \frac{248}{27}$$

d)
$$E(2X + 3Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (2x + 3y)f(x, y)\,dx\,dy = \int_{x=0}^{4}\int_{y=1}^{5} (2x + 3y)\left(\frac{xy}{96}\right)dx\,dy = \frac{47}{3}$$

Otro método

c) Dado que X y Y son independientes, tenemos, empleando los incisos *a*) y *b*),

$$E(XY) = E(X)E(Y) = \left(\frac{8}{3}\right)\left(\frac{31}{9}\right) = \frac{248}{27}$$

d) De acuerdo con el teorema 3-1 y con el teorema 3-2, de las páginas 76-77, junto con los incisos *a*) y *b*),

$$E(2X + 3Y) = 2E(X) + 3E(Y) = 2\left(\frac{8}{3}\right) + 3\left(\frac{31}{9}\right) = \frac{47}{3}$$

3.6. Demostrar el teorema 3-2 de la página 77.

Sea $f(x, y)$ la función de probabilidad conjunta de X y Y, que se supone son discretas. Entonces,

$$E(X + Y) = \sum_x \sum_y (x + y)f(x, y)$$
$$= \sum_x \sum_y xf(x, y) + \sum_x \sum_y yf(x, y)$$
$$= E(X) + E(Y)$$

Si cualquiera de estas variables es continua, la prueba es igual y se desarrolla de la misma manera, pero se sustituyen los correspondientes signos de sumatoria por integraciones. Vemos que el teorema se satisface ya sea que X y Y sean o no independientes.

3.7. Demostrar el teorema 3-3 de la página 77.

Sea $f(x, y)$ la función de probabilidad conjunta de X y Y, que se supone son discretas. Si las variables X y Y son independientes, se tiene que $f(x, y) = f_1(x) f_2(y)$. Por tanto,

$$\begin{aligned} E(XY) &= \sum_x \sum_y xyf(x, y) = \sum_x \sum_y xyf_1(x)f_2(y) \\ &= \sum_x \left[xf_1(x) \sum_y yf_2(y) \right] \\ &= \sum_x [(xf_1(x)E(y)] \\ &= E(X)E(Y) \end{aligned}$$

Si cualquiera de estas variables es continua, la prueba se desarrolla de la misma manera que el problema anterior, pero se sustituye las sumatorias correspondientes por integrales. Vemos que la validez de este teorema depende de que $f(x, y)$ pueda expresarse como una función de x multiplicada por una función de y, para toda x y y, es decir, de que X y Y sean independientes. Para variables dependientes, en general, este teorema no es válido.

VARIANZA Y DESVIACIÓN ESTÁNDAR

3.8. Encontrar *a*) la varianza, *b*) la desviación estándar de la suma de puntos que se obtiene al lanzar un par de dados no cargados.

a) De acuerdo con el problema 3.2, tenemos $E(X) = E(Y) = 7/2$. Además,

$$E(X^2) = E(Y^2) = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + \cdots + 6^2\left(\frac{1}{6}\right) = \frac{91}{6}$$

Entonces, de acuerdo con el teorema 3-4,

$$\text{Var}(X) = \text{Var}(Y) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

y, puesto que X y Y son independientes, el teorema 3-7 proporciona

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) = \frac{35}{6}$$

b)
$$\sigma_{X+Y} = \sqrt{\text{Var}(X + Y)} = \sqrt{\frac{35}{6}}$$

3.9. Encontrar *a*) la varianza, *b*) la desviación estándar de la variable aleatoria del problema 3.4.

a) Como en el problema 3.4, la media de X es $\mu = E(X) = \frac{1}{2}$. Entonces la varianza es

$$\text{Var}(X) = E[(X - \mu)^2] = E\left[\left(X - \frac{1}{2}\right)^2\right] = \int_{-\infty}^{\infty}\left(x - \frac{1}{2}\right)^2 f(x)\,dx$$

$$= \int_0^{\infty}\left(x - \frac{1}{2}\right)^2 (2e^{-2x})\,dx = \frac{1}{4}$$

Otro método

De acuerdo con el teorema 3-4,

$$\text{Var}(X) = E[(X - \mu)^2] = E(X^2) - [E(X)]^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

b)
$$\sigma = \sqrt{\text{Var}(X)} = \sqrt{\frac{1}{4}} = \frac{1}{2}$$

3.10. Demostrar el teorema 3-4 de la página 78.

Tenemos

$$\begin{aligned} E[(X-\mu)^2] &= E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

3.11. Demostrar el teorema 3-6 de la página 78.

$$\begin{aligned} E[(X-a)^2] &= E[\{(X-\mu) + (\mu - a)\}^2] \\ &= E[(X-\mu)^2 + 2(X-\mu)(\mu - a) + (\mu - a)^2] \\ &= E[(X-\mu)^2] + 2(\mu - a)E(X-\mu) + (\mu - a)^2 \\ &= E[(X-\mu)^2] + (\mu - a)^2 \end{aligned}$$

ya que $E(X - \mu) = E(X) - \mu = 0$. A partir de lo anterior vemos que el valor mínimo de $E[(X-a)^2]$ se presenta cuando $(\mu - a)^2 = 0$, es decir, cuando $a = \mu$.

3.12. Si $X^* = (X - \mu)/\sigma$ es una variable aleatoria estandarizada, demostrar que *a*) $E(X^*) = 0$, *b*) $\text{Var}(X^*) = 1$,

a)
$$E(X^*) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}[E(X-\mu)] = \frac{1}{\sigma}[E(X) - \mu] = 0$$

dado que $E(X) = \mu$.

b)
$$\text{Var}(X^*) = \text{Var}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}E[(X-\mu)^2] = 1$$

Usando el teorema 3-5 de la página 78 y el hecho de que $E[(X-\mu)^2] = \sigma^2$.

3.13. Demostrar el teorema 3-7 de la página 78.

$$\begin{aligned} \text{Var}(X+Y) &= E[\{(X+Y) - (\mu_X + \mu_Y)\}^2] \\ &= E[\{(X-\mu_X) + (Y-\mu_Y)\}^2] \\ &= E[(X-\mu_X)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2] \\ &= E[(X-\mu_X)^2] + 2E[(X-\mu_X)(Y-\mu_Y)] + E[(Y-\mu_Y)^2] \\ &= \text{Var}(X) + \text{Var}(Y) \end{aligned}$$

usamos el hecho de que

$$E[(X-\mu_X)(Y-\mu_Y)] = E(X-\mu_X)E(Y-\mu_Y) = 0$$

ya que X y Y y, por tanto, $X - \mu_X$ y $Y - \mu_Y$, son independientes. En la prueba de (19), de la página 78, se sustituye Y por $-Y$ y se usa el teorema 3-5.

MOMENTOS Y FUNCIONES GENERADORAS DE MOMENTOS

3.14. Demostrar el resultado (26) de la página 79.

$$\begin{aligned} \mu_r &= E[(X-\mu)^r] \\ &= E\left[X^r - \binom{r}{1}X^{r-1}\mu + \cdots + (-1)^j\binom{r}{j}X^{r-j}\mu^j \right. \\ &\qquad \left. + \cdots + (-1)^{r-1}\binom{r}{r-1}X\mu^{r-1} + (-1)^r\mu^r\right] \end{aligned}$$

$$= E(X^r) - \binom{r}{1}E(X^{r-1})\mu + \cdots + (-1)^j\binom{r}{j}E(X^{r-j})\mu^j$$

$$+ \cdots + (-1)^{r-1}\binom{r}{r-1}E(X)\mu^{r-1} + (-1)^r\mu^r$$

$$= \mu'_r - \binom{r}{1}\mu'_{r-1}\mu + \cdots + (-1)^j\binom{r}{j}\mu'_{r-j}\mu^j$$

$$+ \cdots + (-1)^{r-1}r\mu^r + (-1)^{-r}\mu^r$$

donde los últimos dos términos pueden combinarse para dar $(-1)^{r-1}(r-1)\mu^r$.

3.15. Demostrar *a*) el resultado (31), *b*) el resultado (32) de la página 79.

a) Cuando se aplica a e^u la expansión en series de potencias (3., apéndice A), tenemos

$$M_X(t) = E(e^{tX}) = E\left(1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots\right)$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \cdots$$

$$= 1 + \mu t + \mu'_2\frac{t^2}{2!} + \mu'_3\frac{t^3}{3!} + \cdots$$

b) Ello es consecuencia inmediata del hecho conocido del cálculo, de que la serie de Taylor de $f(t)$ alrededor de $t = a$ es

$$f(t) = \sum_{n=0}^{\infty} c_n(t-a)^n$$

por lo que

$$c_n = \frac{1}{n!}\frac{d^n}{dt^n}f(t)\bigg|_{t=a}$$

3.16. Demostrar el teorema 3-9 de la página 80.

Dado que X y Y son independientes, cualquier función de X y cualquier función de Y son independientes. Por tanto,

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E(e^{tX}e^{tY}) = E(e^{tX})E(e^{tY}) = M_X(t)M_Y(t)$$

3.17. La variable aleatoria X puede tomar los valores 1 y -1, cada uno con probabilidad $\frac{1}{2}$. Encontrar *a*) la función generadora de momentos, *b*) los primeros cuatro momentos alrededor del origen.

a)
$$E(e^{tX}) = e^{t(1)}\left(\frac{1}{2}\right) + e^{t(-1)}\left(\frac{1}{2}\right) = \frac{1}{2}(e^t + e^{-t})$$

b) Tenemos
$$e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots$$

$$e^{-t} = 1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \frac{t^4}{4!} - \cdots$$

En consecuencia, (1)
$$\frac{1}{2}(e^t + e^{-t}) = 1 + \frac{t^2}{2!} + \frac{t^4}{4!} + \cdots$$

Pero (2)
$$M_X(t) = 1 + \mu t + \mu'_2\frac{t^2}{2!} + \mu'_3\frac{t^3}{3!} + \mu'_4\frac{t^4}{4!} + \cdots$$

Entonces, comparando (1) y (2), se tiene

$$\mu = 0, \qquad \mu'_2 = 1, \qquad \mu'_3 = 0, \qquad \mu'_4 = 1, \ldots$$

Los momentos impares son cero y los momentos pares valen uno.

3.18. Una variable aleatoria X tiene la función de densidad dada por

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Encontrar *a*) la función generadora de momentos, *b*) los primeros cuatro momentos alrededor del origen.

a)
$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \int_0^{\infty} e^{tx}(2e^{-2x})\,dx = 2\int_0^{\infty} e^{(t-2)x}dx$$

$$= \frac{2e^{(t-2)x}}{t-2}\bigg|_0^{\infty} = \frac{2}{2-t}, \quad \text{suponiendo que } t < 2$$

b) Si $|t| < 2$ se tiene

$$\frac{2}{2-t} = \frac{1}{1-t/2} = 1 + \frac{t}{2} + \frac{t^2}{4} + \frac{t^3}{8} + \frac{t^4}{16} + \cdots$$

Pero
$$M(t) = 1 + \mu t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \mu_4' \frac{t^4}{4!} + \cdots$$

Por tanto, comparando términos $\mu = \frac{1}{2}$, $\mu_2' = \frac{1}{2}$, $\mu_3' = \frac{3}{4}$, $\mu_4' = \frac{3}{2}$.

3.19. Determinar los primeros cuatro momentos *a*) alrededor del origen, *b*) alrededor de la media, de una variable aleatoria X cuya función de densidad es

$$f(x) = \begin{cases} 4x(9 - x^2)/81 & 0 \leq x \leq 3 \\ 0 & \text{si no es así} \end{cases}$$

a)
$$\mu_1' = E(X) = \frac{4}{81}\int_0^3 x^2(9 - x^2)\,dx = \frac{8}{5} = \mu$$

$$\mu_2' = E(X^2) = \frac{4}{81}\int_0^3 x^3(9 - x^2)\,dx = 3$$

$$\mu_3' = E(X^3) = \frac{4}{81}\int_0^3 x^4(9 - x^2)\,dx = \frac{216}{35}$$

$$\mu_4' = E(X^4) = \frac{4}{81}\int_0^3 x^5(9 - x^2)\,dx = \frac{27}{2}$$

b) Con base en el resultado (27) de la página 79, tenemos

$$\mu_1 = 0$$

$$\mu_2 = 3 - \left(\frac{8}{5}\right)^2 = \frac{11}{25} = \sigma^2$$

$$\mu_3 = \frac{216}{35} - 3(3)\left(\frac{8}{5}\right) + 2\left(\frac{8}{5}\right)^3 = -\frac{32}{875}$$

$$\mu_4 = \frac{27}{2} - 4\left(\frac{216}{35}\right)\left(\frac{8}{5}\right) + 6(3)\left(\frac{8}{5}\right)^2 - 3\left(\frac{8}{5}\right)^4 = \frac{3\,693}{8\,750}$$

FUNCIONES CARACTERÍSTICAS

3.20. Calcular la función característica de la variable aleatoria X del problema 3.17.

La función característica está dada por

$$E(e^{i\omega X}) = e^{i\omega(1)}\left(\frac{1}{2}\right) + e^{i\omega(-1)}\left(\frac{1}{2}\right) = \frac{1}{2}(e^{i\omega} + e^{-i\omega}) = \cos\omega$$

empleando las fórmulas de Euler,

$$e^{i\theta} = \cos\theta + i\,\text{sen}\,\theta \qquad e^{-i\theta} = \cos\theta - i\,\text{sen}\,\theta$$

con $\theta = \omega$. Este resultado también lo obtenemos del problema 3.17*a*) haciendo $t = i\omega$.

3.21. Determinar la función característica de la variable aleatoria X cuya función de densidad es

$$f(x) = \begin{cases} 1/2a & |x| < a \\ 0 & \text{si no es así} \end{cases}$$

La función característica está dada por

$$E(e^{i\omega X}) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\,dx = \frac{1}{2a}\int_{-a}^{a} e^{i\omega x}\,dx$$

$$= \frac{1}{2a}\frac{e^{i\omega x}}{i\omega}\bigg|_{-a}^{a} = \frac{e^{ia\omega} - e^{-ia\omega}}{2ia\omega} = \frac{\text{sen}\,a\omega}{a\omega}$$

usando las fórmulas de Euler (vea el problema 3.20) con $\theta = a\omega$.

3.22. Encontrar la función característica de la variable aleatoria X cuya función de densidad es $f(x) = ce^{-a|x|}$, $-\infty < x < \infty$, donde $a > 0$ y c es una constante adecuada.

Como $f(x)$ es una función de densidad, tenemos

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

de manera que

$$c\int_{-\infty}^{\infty} e^{-a|x|}dx = c\left[\int_{-\infty}^{0} e^{-a(-x)}dx + \int_{0}^{\infty} e^{-a(x)}dx\right]$$

$$= c\frac{e^{ax}}{a}\bigg|_{-\infty}^{0} + c\frac{e^{-ax}}{-a}\bigg|_{0}^{\infty} = \frac{2c}{a} = 1$$

Entonces, $c = a/2$. La función característica es, por tanto, la dada por

$$E(e^{i\omega X}) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\,dx$$

$$= \frac{a}{2}\left[\int_{-\infty}^{0} e^{i\omega x}e^{-a(-x)}dx + \int_{0}^{\infty} e^{i\omega x}e^{-a(x)}dx\right]$$

$$= \frac{a}{2}\left[\int_{-\infty}^{0} e^{(a+i\omega)x}dx + \int_{0}^{\infty} e^{-(a-i\omega)x}dx\right]$$

$$= \frac{a}{2}\frac{e^{(a+i\omega)x}}{a+i\omega}\bigg|_{-\infty}^{0} + a\frac{e^{-(a-i\omega)x}}{-(a-i\omega)}\bigg|_{0}^{\infty}$$

$$= \frac{a}{2(a+i\omega)} + \frac{a}{2(a-i\omega)} = \frac{a^2}{a^2+\omega^2}$$

COVARIANZA Y COEFICIENTE DE CORRELACIÓN

3.23. Demostrar el teorema 3-14 de la página 81.

Por definición, la covarianza de X y Y es

$$\begin{aligned}\sigma_{XY} &= \text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y] \\ &= E(XY) - \mu_X E(Y) - \mu_Y E(X) + E(\mu_X\mu_Y) \\ &= E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y \\ &= E(XY) - \mu_X\mu_Y \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

3.24. Demostrar el teorema 3-15 de la página 81.

Si X y Y son independientes, entonces $E(XY) = E(X)E(Y)$. Por tanto, de acuerdo con el problema 3.23,

$$\sigma_{XY} = \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

3.25. Determinar *a*) $E(X)$, *b*) $E(Y)$, *c*) $E(XY)$, *d*) $E(X^2)$, *e*) $E(Y^2)$, *f*) Var (X), *g*) Var (Y), *h*) Cov (X, Y), *i*) ρ, si las variables aleatorias X y Y están definidas como en el problema 2.8 de las páginas 47-48.

a)
$$E(X) = \sum_x \sum_y xf(x, y) = \sum_x x\left[\sum_y f(x, y)\right]$$
$$= (0)(6c) + (1)(14c) + (2)(22c) = 58c = \frac{58}{42} = \frac{29}{21}$$

b)
$$E(Y) = \sum_x \sum_y yf(x, y) = \sum_y y\left[\sum_x f(x, y)\right]$$
$$= (0)(6c) + (1)(9c) + (2)(12c) + (3)(15c) = 78c = \frac{78}{42} = \frac{13}{7}$$

c)
$$E(XY) = \sum_x \sum_y xyf(x, y)$$
$$= (0)(0)(0) + (0)(1)(c) + (0)(2)(2c) + (0)(3)(3c)$$
$$+ (1)(0)(2c) + (1)(1)(3c) + (1)(2)(4c) + (1)(3)(5c)$$
$$+ (2)(0)(4c) + (2)(1)(5c) + (2)(2)(6c) + (2)(3)(7c)$$
$$= 102c = \frac{102}{42} = \frac{17}{7}$$

d)
$$E(X^2) = \sum_x \sum_y x^2 f(x, y) = \sum_x x^2\left[\sum_y f(x, y)\right]$$
$$= (0)^2(6c) + (1)^2(14c) + (2)^2(22c) = 102c = \frac{102}{42} = \frac{17}{7}$$

e)
$$E(Y^2) = \sum_x \sum_y y^2 f(x, y) = \sum_y y^2\left[\sum_x f(x, y)\right]$$
$$= (0)^2(6c) + (1)^2(9c) + (2)^2(12c) + (3)^2(15c) = 192c = \frac{192}{42} = \frac{32}{7}$$

f)
$$\sigma_X^2 = \text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{17}{7} - \left(\frac{29}{21}\right)^2 = \frac{230}{441}$$

g)
$$\sigma_Y^2 = \text{Var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{32}{7} - \left(\frac{13}{7}\right)^2 = \frac{55}{49}$$

h)
$$\sigma_{XY} = \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{17}{7} - \left(\frac{29}{21}\right)\left(\frac{13}{7}\right) = -\frac{20}{147}$$

i)
$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-20/147}{\sqrt{230/441}\sqrt{55/49}} = \frac{-20}{\sqrt{230}\sqrt{55}} = -0.2103 \text{ aprox.}$$

3.26. Repetir el problema 3.25 si las variables aleatorias X y Y están definidas como en el problema 2.33 de las páginas 61-63.

Usando $c = 1/210$, tenemos:

a)
$$E(X) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5} (x)(2x + y)\,dx\,dy = \frac{268}{63}$$

b)
$$E(Y) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5} (y)(2x + y)\,dx\,dy = \frac{170}{63}$$

c)
$$E(XY) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5} (xy)(2x + y)\,dx\,dy = \frac{80}{7}$$

d) $$E(X^2) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5}(x^2)(2x + y)\,dx\,dy = \frac{1\,220}{63}$$

e) $$E(Y^2) = \frac{1}{210}\int_{x=2}^{6}\int_{y=0}^{5}(y^2)(2x + y)\,dx\,dy = \frac{1\,175}{126}$$

f) $$\sigma_X^2 = \mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{1\,220}{63} - \left(\frac{268}{63}\right)^2 = \frac{5\,036}{3\,969}$$

g) $$\sigma_Y^2 = \mathrm{Var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{1\,175}{126} - \left(\frac{170}{63}\right)^2 = \frac{16\,225}{7\,938}$$

h) $$\sigma_{XY} = \mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{80}{7} - \left(\frac{268}{63}\right)\left(\frac{170}{63}\right) = -\frac{200}{3\,969}$$

i) $$\rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{-\,200/3\,969}{\sqrt{5\,036/3\,969}\sqrt{16\,225/7\,938}} = \frac{-200}{\sqrt{2\,518}\sqrt{16\,225}} = -0.03129 \text{ aprox.}$$

ESPERANZA, VARIANZA Y MOMENTOS CONDICIONALES

3.27. En el problema 2.8 de las páginas 47-48, calcular la esperanza condicional de Y dado $X = 2$.

Como en el problema 2.27, página 58, la función de probabilidad condicional de Y dada $X = 2$ es

$$f(y\,|\,2) = \frac{4 + y}{22}$$

En consecuencia, la esperanza condicional de Y dada $X = 2$ es

$$E(Y\,|\,X = 2) = \sum_y y\left(\frac{4 + y}{22}\right)$$

donde se suman todas las y que corresponden a $X = 2$. Ello está dado por

$$E(Y\,|\,X = 2) = (0)\left(\frac{4}{22}\right) + 1\left(\frac{5}{22}\right) + 2\left(\frac{6}{22}\right) + 3\left(\frac{7}{22}\right) = \frac{19}{11}$$

3.28. Determinar la esperanza condicional de *a*) Y dada X, *b*) X dada Y en el problema 2.29, páginas 58-59.

a) $$E(Y\,|\,X = x)\int_{-\infty}^{\infty} y f_2(y\,|\,x)\,dy = \int_0^x y\left(\frac{2y}{x^2}\right)dy = \frac{2x}{3}$$

b) $$E(X\,|\,Y = y) = \int_{-\infty}^{\infty} x f_1(x\,|\,y)\,dx = \int_y^1 x\left(\frac{2x}{1 - y^2}\right)dx$$

$$= \frac{2(1 - y^3)}{3(1 - y^2)} = \frac{2(1 + y + y^2)}{3(1 + y)}$$

3.29. Encontrar la varianza condicional de Y dada X en el problema 2.29, páginas 58-59.

La varianza que se busca (segundo momento alrededor de la media) está dada por

$$E[(Y - \mu_2)^2\,|\,X = x] = \int_{-\infty}^{\infty}(y - \mu_2)^2 f_2(y\,|\,x)\,dy = \int_0^x\left(y - \frac{2x}{3}\right)^2\left(\frac{2y}{x^2}\right)dy = \frac{x^2}{18}$$

donde usamos el hecho de que $\mu_2 = E(Y\,|\,X = x) = 2x/3$ de acuerdo con el problema 3.28*a*).

DESIGUALDAD DE CHEBYSHEV

3.30. Demostrar la desigualdad de Chebyshev.

Se presentará la prueba para variables aleatorias continuas. La demostración para variables aleatorias discretas es similar, sustituyendo solamente las integrales por sumas. Si $f(x)$ es la función de densidad de X, entonces

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)\,dx$$

Como el integrando es no negativo, el valor de la integral sólo puede disminuir cuando el intervalo de integración se reduce. Por tanto,

$$\sigma^2 \geq \int_{|x-\mu| \geq \epsilon} (x - \mu)^2 f(x)\,dx \geq \int_{|x-\mu| \geq \epsilon} \epsilon^2 f(x)\,dx = \epsilon^2 \int_{|x-\mu| \geq \epsilon} f(x)\,dx$$

Pero la última integral es igual a $P(|X - \mu| \geq \epsilon)$. De manera que,

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

3.31. Dada la variable aleatoria del problema 3.18, *a*) determinar $P(|X - \mu| > 1)$. *b*) Utilizar la desigualdad de Chebyshev para obtener la cota superior de $P(|X - \mu| > 1)$ y comparar con el resultado del inciso *a*).

a) De acuerdo con el problema 3.18, $\mu = 1/2$. Entonces,

$$P(|X - \mu| < 1) = P\left(\left|X - \frac{1}{2}\right| < 1\right) = P\left(-\frac{1}{2} < X < \frac{3}{2}\right)$$

$$= \int_0^{3/2} 2e^{-2x}dx = 1 - e^{-3}$$

Por tanto, $$P\left(\left|X - \frac{1}{2}\right| \geq 1\right) = 1 - (1 - e^{-3}) = e^{-3} = 0.04979$$

b) De acuerdo con el problema 3.18, $\sigma^2 = \mu_2' - \mu^2 = 1/4$. La desigualdad de Chebyshev con $\epsilon = 1$ da como resultado

$$P(|X - \mu| \geq 1) \leq \sigma^2 = 0.25$$

Si lo comparamos con *a*), vemos que la cota que proporciona la desigualdad de Chebyshev es aquí bastante inexacta. En la práctica, la desigualdad de Chebyshev se usa para obtener estimaciones cuando no es muy necesario o es imposible obtener valores exactos.

LEY DE LOS GRANDES NÚMEROS

3.32. Demostrar la ley de los grandes números dada en el teorema 3-19 de la página 83.

Tenemos $$E(X_1) = E(X_2) = \cdots = E(X_n) = \mu$$

$$\text{Var}(X_1) = \text{Var}(X_2) = \cdots = \text{Var}(X_n) = \sigma^2$$

En consecuencia, $$E\left(\frac{S_n}{n}\right) = E\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n}[E(X_1) + \cdots + E(X_n)] = \frac{1}{n}(n\mu) = \mu$$

$$\text{Var}(S_n) = \text{Var}(X_1 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) = n\sigma^2$$

de manera que $$\text{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\text{Var}(S_n) = \frac{\sigma^2}{n}$$

donde se ha usado el teorema 3-5 y una extensión del teorema 3-7.

Por tanto, de acuerdo con la desigualdad de Chebyshev con $X = S_n/n$, tenemos

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$

Tomando el límite cuando $n \to \infty$, este resultado se convierte en

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = 0$$

como se buscaba.

OTRAS MEDIDAS DE TENDENCIA CENTRAL

3.33. La función de densidad de una variable aleatoria continua X es

$$f(x) = \begin{cases} 4x(9 - x^2)/81 & 0 \leq x \leq 3 \\ 0 & \text{si no es así} \end{cases}$$

a) Calcular la moda. *b*) Determinar la mediana. *c*) Comparar moda, mediana y media.

a) La moda se obtiene cuando se determina dónde tiene un máximo relativo la densidad $f(x)$. El máximo relativo de $f(x)$ se presenta donde la derivada es cero, es decir,

$$\frac{d}{dx}\left[\frac{4x(9 - x^2)}{81}\right] = \frac{36 - 12x^2}{81} = 0$$

Por consiguiente, $x = \sqrt{3} = 1.73$ aproximadamente, que es la moda buscada. Vemos que esto da el máximo ya que la segunda derivada, $-24x/81$, es negativa para $x = \sqrt{3}$.

b) La mediana es aquel valor a para el que $P(X \leq a) = 1/2$. Ahora, dado que $0 < a < 3$.

$$P(X \leq a) = \frac{4}{81}\int_0^a x(9 - x^2)\,dx = \frac{4}{81}\left(\frac{9a^2}{2} - \frac{a^4}{4}\right)$$

Igualando esta expresión a $1/2$, determinamos que

$$2a^4 - 36a^2 + 81 = 0$$

de donde

$$a^2 = \frac{36 \pm \sqrt{(36)^2 - 4(2)(81)}}{2(2)} = \frac{36 \pm \sqrt{648}}{4} = 9 \pm \frac{9}{2}\sqrt{2}$$

Por tanto, la mediana que buscamos, que debe estar entre 0 y 3, está dada por

$$a^2 = 9 - \frac{9}{2}\sqrt{2}$$

de donde $a = 1.62$, aproximadamente.

c)
$$E(X) = \frac{4}{81}\int_0^3 x^2(9 - x^2)\,dx = \frac{4}{81}\left(3x^3 - \frac{x^5}{5}\right)\Big|_0^3 = 1.60$$

que es casi igual a la mediana. En la figura 3-6 se muestran la moda, la mediana y la media.

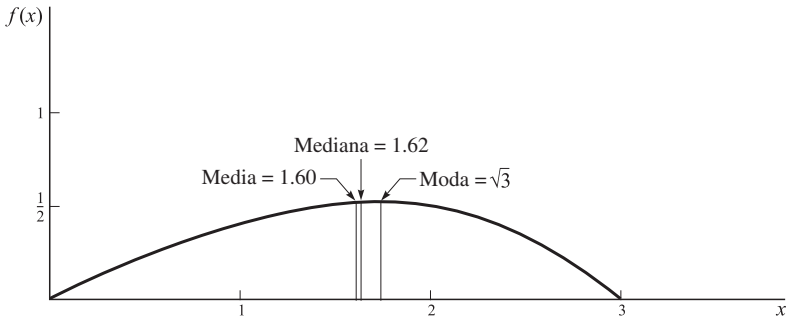


Figura 3-6

3.34. Una variable aleatoria discreta tiene la función de probabilidad $f(x) = 1/2^x$ donde $x = 1, 2, \ldots$ Determinar *a*) la moda, *b*) la mediana y *c*) compare moda, mediana y media.

a) La moda es el valor x que tiene la mayor probabilidad. En este caso $x = 1$, para el que la probabilidad es $1/2$.

b) Si x es cualquier valor entre 1 y 2, $P(X < x) = \frac{1}{2}$ y $P(X > x) = \frac{1}{2}$. Por tanto, *cualquier número* entre 1 y 2 puede emplearse como mediana. Por conveniencia, se elige el punto medio del intervalo, es decir $3/2$.

c) Como encontramos en el problema 3.3, $\mu = 2$. Por tanto, el orden en que se presentan estas tres medidas es precisamente el contrario al que obtuvimos en el problema 3.33.

PERCENTILES

3.35. Determinar los valores correspondientes a los percentiles *a*) décimo, *b*) vigésimo quinto y *c*) septuagésimo quinto de la distribución del problema 3.33.

De acuerdo con el problema 3.33*b*) tenemos

$$P(X \le a) = \frac{4}{81}\left(\frac{9a^2}{2} - \frac{a^4}{4}\right) = \frac{18a^2 - a^4}{81}$$

a) El décimo percentil es el valor de a para el que $P(X \le a) = 0.10$, es decir, la solución de $(18a^2 - a^4)/81 = 0.10$. Empleando el método del problema 3.33 obtenemos $a = 0.68$ aproximadamente.

b) El vigésimo quinto percentil es el valor de a tal que $(18a^2 - a^4)/81 = 0.25$, y encontramos $a = 1.098$, aproximadamente.

c) El septuagésimo quinto percentil es el valor de a tal que $(18a^2 - a^4)/81 = 0.75$, esto es, $a = 2.121$, aproximadamente.

OTRAS MEDIDAS DE DISPERSIÓN

3.36. Determinar, *a*) el rango semiintercuartil, *b*) la desviación media de la distribución del problema 3.33.

a) De acuerdo con el problema 3.35, los valores del vigésimo quinto y del septuagésimo quinto percentiles son 1.098 y 2.121, respectivamente. Por tanto,

$$\text{Rangos semiintercuartil} = \frac{2.121 - 1.098}{2} = 0.51 \text{ aproximadamente}$$

b) De acuerdo con el problema 3.33, la media es $\mu = 1.60 = 8/5$. Entonces,

$$\text{Desviación media} = \text{D.M.} = E(|X - \mu|) = \int_{-\infty}^{\infty} |x - \mu| f(x)\,dx$$

$$= \int_0^3 \left| x - \frac{8}{5} \right| \left[\frac{4x}{81}(9 - x^2)\right] dx$$

$$= \int_0^{8/5} \left(\frac{8}{5} - x\right)\left[\frac{4x}{81}(9 - x^2)\right] dx + \int_{8/5}^3 \left(x - \frac{8}{5}\right)\left[\frac{4x}{81}(9 - x^2)\right] dx$$

$$= 0.555 \text{ aproximadamente}$$

SESGO Y CURTOSIS

3.37. Encontrar el coeficiente *a*) de sesgo, *b*) de curtosis de la distribución del problema 3.19.

De acuerdo con el problema 3.19*b*) tenemos

$$\sigma^2 = \frac{11}{25} \qquad \mu_3 = -\frac{32}{875} \qquad \mu_4 = \frac{3\,693}{8\,750}$$

a) Coeficiente de sesgo $= \alpha_3 = \dfrac{\mu_3}{\sigma^3} = -0.1253$

b) Coeficiente de curtosis $= \alpha_4 = \dfrac{\mu_4}{\sigma^4} = 2.172$

Se deduce que existe un sesgo moderado a la izquierda, como lo indica la figura 3-6. Esta distribución también tiene un pico algo menos puntiagudo que la distribución normal, cuya curtosis es de 3.

PROBLEMAS DIVERSOS

3.38. Si $M(t)$ es la función generadora de momentos de una variable aleatoria X, demostrar que la media es $\mu = M'(0)$ y la varianza es $\sigma^2 = M''(0) - [M'(0)]^2$.

De acuerdo con (32), página 79, haciendo $r = 1$ y $r = 2$,

$$\mu_1' = M'(0) \qquad \mu_2' = M''(0)$$

Por consiguiente, de acuerdo con (27)

$$\mu = M'(0) \qquad \mu_2 = \sigma^2 = M''(0) - [M'(0)]^2$$

3.39. Sea X una variable aleatoria que tome los valores $x_k = k$ con probabilidades p_k, donde $k = \pm 1, \ldots, \pm n$. *a*) Encontrar la función característica $\phi(\omega)$ de X, *b*) calcular p_k en términos de $\phi(\omega)$.

a) La función característica es

$$\phi(\omega) = E(e^{i\omega X}) = \sum_{k=-n}^{n} e^{i\omega x_k} p_k = \sum_{k=-n}^{n} p_k e^{ik\omega}$$

b) Multiplicamos ambos lados de la expresión del inciso *a*) por $e^{-ij\omega}$ e integramos respecto a ω de 0 a 2π. Entonces,

$$\int_{\omega=0}^{2\pi} e^{-ij\omega}\phi(\omega)\,d\omega = \sum_{k=-n}^{n} p_k \int_{\omega=0}^{2\pi} e^{i(k-j)\omega}\,d\omega = 2\pi p_j$$

ya que

$$\int_{\omega=0}^{2\pi} e^{i(k-j)\omega}\,d\omega = \begin{cases} \left.\dfrac{e^{i(k-j)\omega}}{i(k-j)}\right|_0^{2\pi} = 0 & k \neq j \\ 2\pi & k = j \end{cases}$$

Por tanto,

$$p_j = \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ij\omega}\phi(\omega)\,d\omega$$

o bien, sustituyendo k en vez de j,

$$p_k = \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ik\omega}\phi(\omega)\,d\omega$$

A $\sum_{k=-n}^{n} p_k e^{ik\omega}$ (donde n teóricamente puede ser infinito) se le suele conocer como la *serie de Fourier* de $\phi(\omega)$ y p_k como los *coeficientes de Fourier*. En el caso de una variable aleatoria continua, la serie de Fourier se sustituye por la integral de Fourier (vea la página 81).

3.40. Use el problema 3.39 para obtener la distribución de probabilidad de una variable aleatoria X cuya función característica es $\phi(\omega) = \cos\omega$.

De acuerdo con el problema 3.39

$$\begin{aligned} p_k &= \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ik\omega}\cos\omega\,d\omega \\ &= \frac{1}{2\pi}\int_{\omega=0}^{2\pi} e^{-ik\omega}\left[\frac{e^{i\omega} + e^{-i\omega}}{2}\right]d\omega \\ &= \frac{1}{4\pi}\int_{\omega=0}^{2\pi} e^{i(1-k)\omega}\,d\omega + \frac{1}{4\pi}\int_{\omega=0}^{2\pi} e^{-i(1+k)\omega}\,d\omega \end{aligned}$$

Si $k = 1$, se determina $p_1 = \frac{1}{2}$; si $k = -1$, encontramos $p_{-1} = \frac{1}{2}$. Para todos los demás valores de k, tenemos $p_k = 0$. Por tanto, la variable aleatoria está dada por

$$X = \begin{cases} 1 & \text{probabilidad } 1/2 \\ -1 & \text{probabilidad } 1/2 \end{cases}$$

Como verificación, vea el problema 3.20.

3.41. Encontrar el coeficiente *a*) de sesgo, *b*) de curtosis de la distribución definida por la *curva normal*, cuya densidad es

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad -\infty < x < \infty$$

a) El comportamiento de esta distribución es el que se muestra en la figura 3-7. Por simetría $\mu_1' = \mu = 0$ y $\mu_3' = 0$. Por tanto, el coeficiente de sesgo es cero.

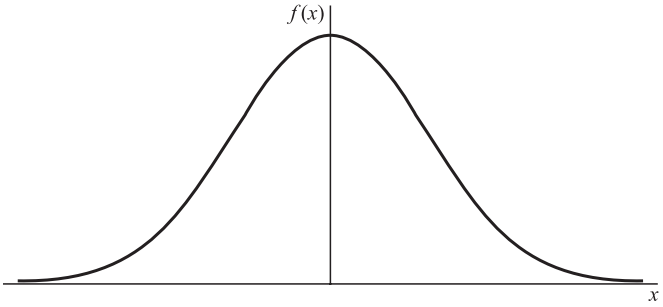


Figura 3-7

b) Tenemos

$$\mu_2' = E(X^2) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} x^2 e^{-x^2/2}\,dx$$
$$= \frac{2}{\sqrt{\pi}}\int_{0}^{\infty} v^{1/2} e^{-v}\,dv$$
$$= \frac{2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) = \frac{2}{\sqrt{\pi}}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = 1$$

donde hemos hecho la transformación $x^2/2 = v$ y usado las propiedades de la función gama dadas en (2) y (5) del apéndice A. De manera similar tenemos

$$\mu_4' = E(X^4) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^4 e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} x^4 e^{-x^2/2}\,dx$$
$$= \frac{4}{\sqrt{\pi}}\int_{0}^{\infty} v^{3/2} e^{-v}\,dv$$
$$= \frac{4}{\sqrt{\pi}}\Gamma\left(\frac{5}{2}\right) = \frac{4}{\sqrt{\pi}}\cdot\frac{3}{2}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = 3$$

Ahora,

$$\sigma^2 = E[(X-\mu)^2] = E(X)^2 = \mu_2' = 1$$
$$\mu_4 = E[(X-\mu)^4] = E(X^4) = \mu_4' = 3$$

Por tanto, el coeficiente de curtosis es

$$\frac{\mu_4}{\sigma^4} = 3$$

3.42. Demostrar que si $-1 \le \rho \le 1$ (vea página 82).

Para toda constante real c, se tiene

$$E[\{Y - \mu_Y - c(X-\mu)\}^2] \ge 0$$

Ahora, el lado izquierdo puede expresarse como

$$E[(Y-\mu_Y)^2] + c^2E[(X-\mu_X)^2] - 2cE[(X-\mu_X)(Y-\mu_Y)] = \sigma_Y^2 + c^2\sigma_X^2 - 2c\sigma_{XY}$$
$$= \sigma_Y^2 + \sigma_X^2\left(c^2 - \frac{2c\sigma_{XY}}{\sigma_X^2}\right)$$
$$= \sigma_Y^2 + \sigma_X^2\left(c^2 - \frac{\sigma_{XY}}{\sigma_X^2}\right)^2 - \frac{\sigma_{XY}^2}{\sigma_X^2}$$
$$= \frac{\sigma_X^2\sigma_Y^2 - \sigma_{XY}^2}{\sigma_X^2} + \sigma_X^2\left(c - \frac{\sigma_{XY}}{\sigma_X^2}\right)^2$$

Para que la última cantidad sea mayor o igual a cero para cada valor de c, debemos tener:

$$\sigma_X^2\sigma_Y^2 - \sigma_{XY}^2 \geq 0 \quad \text{o bien} \quad \frac{\sigma_{XY}^2}{\sigma_X^2\sigma_Y^2} \leq 1$$

lo que es equivalente a $\rho^2 \leq 1$ o bien $-1 \leq \rho \leq 1$.

PROBLEMAS SUPLEMENTARIOS

ESPERANZA DE VARIABLES ALEATORIAS

3.43. Una variable aleatoria X está definida por $X = \begin{cases} -2 & \text{prob. } 1/3 \\ 3 & \text{prob. } 1/2. \\ 1 & \text{prob. } 1/6 \end{cases}$ Encuentre *a*) $E(X)$, *b*) $E(2X + 5)$, *c*) $E(X^2)$.

3.44. Sea X una variable aleatoria definida mediante la función de densidad $f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{si no es así} \end{cases}$.

Encuentre *a*) $E(X)$, *b*) $E(3X - 2)$, *c*) $E(X^2)$

3.45. La función de densidad de una variable aleatoria X es $f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{si no es así} \end{cases}$.

Encuentre *a*) $E(X)$, *b*) $E(X^2)$, *c*) $E[(X - 1)^2]$.

3.46. ¿Cuál es el número esperado de puntos que se obtendrá en tres lanzamientos sucesivos de un dado no cargado? ¿Parece razonable su respuesta? Explique.

3.47. Una variable aleatoria X tiene la función de densidad $f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$. Encuentre $E(e^{2X/3})$.

3.48. Sean X y Y variables aleatorias independientes que tienen cada una la función de densidad

$$f(u) = \begin{cases} 2e^{-2u} & u \geq 0 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) $E(X + Y)$, *b*) $E(X^2 + Y^2)$, *c*) $E(XY)$.

3.49. En el problema 3.48, ¿*a*) $E(X + Y) = E(X) + E(Y)$, *b*) $E(XY) = E(X)E(Y)$? Explique su respuesta.

3.50. Sean X y Y variables aleatorias que tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} \frac{3}{5}x(x + y) & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) $E(X)$, *b*) $E(Y)$, *c*) $E(X + Y)$, *d*) $E(XY)$.

3.51. En el problema 3.50, ¿*a*) $E(X + Y) = E(X) + E(Y)$, *b*) $E(XY) = E(X)E(Y)$? Explique su respuesta.

3.52. Sean X y Y variables aleatorias que tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} 4xy & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) $E(X)$, *b*) $E(Y)$, *c*) $E(X + \mathrm{Y})$, *d*) $E(XY)$.

3.53. En el problema 3.52, ¿*a*) $E(X + Y) = E(X) + E(Y)$, *b*) $E(XY) = E(X)E(Y)$? Explique su respuesta.

3.54. Sea $f(x, y) = \begin{cases} \frac{1}{4}(2x + y) & 0 \le x \le 1, 0 \le y \le 2 \\ 0 & \text{si no es así} \end{cases}$. Encuentre *a*) $E(X)$, *b*) $E(Y)$, *c*) $E(X^2)$, *d*) $E(Y^2)$, *e*) $E(X + Y)$, *f*) $E(XY)$.

3.55. Sean X y Y variables aleatorias independientes tales que

$$X = \begin{cases} 1 & \text{prob. } 1/3 \\ 0 & \text{prob. } 2/3 \end{cases} \qquad Y = \begin{cases} 2 & \text{prob. } 3/4 \\ -3 & \text{prob. } 1/4 \end{cases}$$

Encuentre *a*) $E(3X + 2Y)$, *b*) $E(2X^2 - Y^2)$, *c*) $E(XY)$, *d*) $E(X^2Y)$.

3.56. Sean $X_1, X_2, \ldots, X_n$ n variables aleatorias distribuidas de manera idéntica tales que

$$X_k = \begin{cases} 1 & \text{prob. } 1/2 \\ 2 & \text{prob. } 1/3 \\ -1 & \text{prob. } 1/6 \end{cases}$$

Encuentre *a*) $E(X_1 + X_2 + \cdots + X_n)$, *b*) $E(X_1^2 + X_2^2 + \cdots + X_n^2)$.

VARIANZA Y DESVIACIÓN ESTÁNDAR

3.57. Calcule *a*) la varianza, *b*) la desviación estándar del número de puntos que se obtendrá en un solo lanzamiento de un dado no cargado.

3.58. Sea X una variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} 1/4 & -2 \le x \le 2 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) $\text{Var}(X)$, *b*) σ_X.

3.59. Sea X una variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} e^{-x} & x \ge 0 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) $\text{Var}(X)$, *b*) σ_X.

3.60. Determine la varianza y la desviación estándar de la variable aleatoria X *a*) del problema 3.43, *b*) del problema 3.44.

3.61. Una variable aleatoria X tiene $E(X) = 2$, $E(X^2) = 8$. Determine *a*) $\text{Var}(X)$, *b*) σ_X.

3.62. Si una variable aleatoria X es tal que $E[(X - 1)^2] = 10$, $E[(X - 2)^2] = 6$ calcule *a*) $E(X)$, *b*) $\text{Var}(X)$, *c*) σ_X.

MOMENTOS Y FUNCIONES GENERADORAS DE MOMENTOS

3.63. Encuentre *a*) la función generadora de momentos de la variable aleatoria

$$X = \begin{cases} 1/2 & \text{prob. } 1/2 \\ -1/2 & \text{prob. } 1/2 \end{cases}$$

y *b*) los primeros cuatro momentos alrededor del origen.

3.64. *a*) Calcule la función generadora de momentos de la variable aleatoria X cuya función de densidad es

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & \text{si no es así} \end{cases}$$

b) Use la función generadora del inciso *a*) para determinar los primeros cuatro momentos alrededor del origen.

3.65. Calcule los primeros cuatro momentos alrededor de la media en *a*) el problema 3.43, *b*) el problema 3.44.

3.66. *a*) Determine la función generadora de momentos de una variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} e^{-x} & x \ge 0 \\ 0 & \text{si no es así} \end{cases}$$

y *b*) determine los primeros cuatro momentos alrededor del origen.

3.67. Calcule los primeros cuatro momentos alrededor de la media en el problema 3.66.

3.68. Si X tiene la función de densidad $f(x) = \begin{cases} 1/(b-a) & a \le x \le b \\ 0 & \text{si no es así} \end{cases}$. Encuentre el k-ésimo momento alrededor *a*) del origen, *b*) de la media.

3.69. Si $M(t)$ es la función generadora de momentos de la variable aleatoria X, demuestre que los momentos 3o. y 4o. alrededor de la media están dados por

$$\mu_3 = M'''(0) - 3M''(0)M'(0) + 2[M'(0)]^3$$
$$\mu_4 = M^{(\text{iv})}(0) - 4M'''(0)M'(0) + 6M''(0)[M'(0)]^2 - 3[M'(0)]^4$$

FUNCIONES CARACTERÍSTICAS

3.70. Determine la función característica de la variable aleatoria $X = \begin{cases} a & \text{prob. } p \\ b & \text{prob. } q = 1 - p \end{cases}$.

3.71. Calcule la función característica de la variable aleatoria X cuya función de densidad es

$$f(x) = \begin{cases} 1/2a & |x| \le a \\ 0 & \text{si no es así} \end{cases}$$

3.72. Encuentre la función característica de una variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & \text{si no es así} \end{cases}$$

3.73. Sean $X_k = \begin{cases} 1 & \text{prob. } 1/2 \\ -1 & \text{prob. } 1/2 \end{cases}$ variables aleatorias independientes ($k = 1, 2, \ldots, n$). Demuestre que la función característica de la variable aleatoria

$$\frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}}$$

es $[\cos(\omega/\sqrt{n})]^n$.

3.74. Demuestre que cuando $n \to \infty$, la función característica del problema 3.73 tiende a $e^{-\omega^2/2}$. (*Sugerencia:* Utilice los logaritmos de la función característica y la regla de L'Hopital.)

COVARIANZA Y COEFICIENTE DE CORRELACIÓN

3.75. Sean X y Y variables aleatorias cuya función de densidad conjunta es

$$f(x, y) = \begin{cases} x + y & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) Var(X), *b*) Var(Y), *c*) σ_X, *d*) σ_Y, *e*) σ_{XY}, *f*) ρ.

3.76. Repita el problema 3.75 con la función de densidad conjunta $f(x, y) = \begin{cases} e^{-(x+y)} & x \ge 0, y \ge 0 \\ 0 & \text{si no es así} \end{cases}$.

3.77. Determine *a*) Var(X), *b*) Var(Y), *c*) σ_X, *d*) σ_Y, *e*) σ_{XY}, *f*) ρ, de las variables aleatorias del problema 2.56.

3.78. Repita el problema 3.77 con las variables aleatorias del problema 2.94.

3.79. Encuentre *a*) la covarianza, *b*) el coeficiente de correlación de dos variables aleatorias X y Y si $E(X) = 2$, $E(Y) = 3$, $E(XY) = 10$, $E(X^2) = 9$, $E(Y^2) = 16$.

3.80. El coeficiente de correlación de dos variables aleatorias X y Y es $-\frac{1}{4}$ y sus varianzas son 3 y 5. Calcule la covarianza.

ESPERANZA, VARIANZA Y MOMENTOS CONDICIONALES

3.81. Si X y Y tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} x + y & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre la esperanza condicional de *a*) Y dado X, *b*) X dado Y.

3.82. Repita el problema 3.81 si $f(x, y) = \begin{cases} 2e^{-(x+2y)} & x \ge 0, y \ge 0 \\ 0 & \text{si no es así} \end{cases}$

3.83. X y Y tienen la función de probabilidad conjunta dada en la tabla 2-9 de la página 71. Determine la esperanza condicional de *a*) Y dada X, *b*) X dada Y.

3.84. Encuentre la varianza condicional de *a*) Y dada X, *b*) X dada Y para la distribución del problema 3.81.

3.85. Repita el problema 3.84 dada la distribución del problema 3.82.

3.86. Repita el problema 3.84 dada la distribución del problema 2.94.

DESIGUALDAD DE CHEBYSHEV

3.87. Una variable aleatoria X tiene media 3 y varianza 2. Use la desigualdad de Chebyshev para obtener la cota superior de *a*) $P(|X - 3| \ge 2)$, *b*) $P(|X - 3| \ge 1)$.

3.88. Demuestre la desigualdad de Chebyshev de una variable discreta X. (*Sugerencia:* Vea el problema 3.30.)

3.89. Una variable aleatoria X tiene como función de densidad $f(x) = \frac{1}{2}e^{-|x|}$, $-\infty < x < \infty$. *a*) Determine $P(|X - \mu| > 2)$. *b*) Use la desigualdad de Chebyshev para obtener la cota superior de $P(|X - \mu| > 2)$ y compare con el resultado del inciso *a*).

LEY DE LOS GRANDES NÚMEROS

3.90. Demuestre que la ley (débil) de los grandes números puede enunciarse como

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) = 1$$

e interprete su respuesta.

3.91. Sean X_k $(k = 1, \ldots, n)$ n variables aleatorias independientes tales que

$$X_k = \begin{cases} 1 & \text{prob. } p \\ 0 & \text{prob. } q = 1 - p \end{cases}$$

a) Si X_k se interpreta como el número de caras al k-ésimo lanzamiento de una moneda, ¿qué interpretación puede dársele a $S_n = X_1 + \cdots + X_n$?

b) Demuestre que en este caso la ley de los grandes números se reduce a

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) = 0$$

e interprete este resultado.

OTRAS MEDIDAS DE TENDENCIA CENTRAL

3.92. Encuentre *a*) la moda, *b*) la mediana de una variable aleatoria X cuya función de densidad es

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{si no es así} \end{cases}$$

y *c*) compárelas con la media.

3.93. Repita el problema 3.100 si la función de densidad es

$$f(x) = \begin{cases} 4x(1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

3.94. Determine *a*) la mediana, *b*) la moda de una variable aleatoria X definida por

$$X = \begin{cases} 2 & \text{prob. } 1/3 \\ -1 & \text{prob. } 2/3 \end{cases}$$

y *c*) compárelas con la media.

3.95. Calcule *a*) la mediana, *b*) la moda del conjunto de números 1, 3, 2, 1, 5, 6, 3, 3 y *c*) compárelas con la media.

PERCENTILES

3.96. Encuentre los valores correspondientes al *a*) vigésimo quinto, y al *b*) septuagésimo quinto percentiles de la variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} 2(1 - x) & 0 \leq x \leq 1 \\ 0 & \text{si no es así} \end{cases}$$

3.97. Encuentre los valores correspondientes a *a*) el décimo, *b*) vigésimo quinto, *c*) septuagésimo quinto y *d*) nonagésimo percentiles de la variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} c(x - x^3) & 0 < x < 1 \\ 0 & \text{si no es así} \end{cases}$$

donde c es una constante apropiada.

OTRAS MEDIDAS DE DISPERSIÓN

3.98. Calcule *a*) el rango semiintercuartil, *b*) la desviación media de la variable aleatoria del problema 3.96.

3.99. Repita el problema 3.98 con la variable aleatoria del problema 3.97.

3.100. Encuentre la desviación media de la variable aleatoria X en cada uno de los casos siguientes.

$$a)\ f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{si no es así} \end{cases} \qquad b)\ f(x) = \frac{1}{\pi(1 + x^2)}, \quad -\infty < x < \infty.$$

3.101. Obtenga la probabilidad de que la variable aleatoria X difiera de su media en más del rango semiintercuartil en el caso *a*) del problema 3.96, *b*) del problema 3.100*a*).

SESGO Y CURTOSIS

3.102. Encuentre el coeficiente *a*) de sesgo, *b*) de curtosis de la distribución del problema 3.100*a*).

3.103. Si

$$f(x) = \begin{cases} c\left(1 - \dfrac{|x|}{a}\right) & |x| \leq a \\ 0 & |x| > a \end{cases}$$

donde c es una constante adecuada, es la función de densidad de X, encuentre el coeficiente *a*) de sesgo, *b*) de curtosis.

3.104. Encuentre el coeficiente *a*) de sesgo, *b*) de curtosis de la distribución cuya función de densidad es

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

PROBLEMAS DIVERSOS

3.105. Sea X una variable aleatoria que puede tomar los valores 2, 1 y 3 con probabilidades $1/3$, $1/6$ y $1/2$, respectivamente. Encuentre *a*) la media, *b*) la varianza, *c*) la función generadora de momentos, *d*) la función característica, *e*) el tercer momento alrededor de la media.

3.106. Repita el problema 3.105 si X tiene la función de densidad

$$f(x) = \begin{cases} c(1 - x) & 0 < x < 1 \\ 0 & \text{si no es así} \end{cases}$$

donde c es una constante adecuada.

3.107. Se lanzan sucesivamente tres dados, que se supone no están cargados. Encuentre *a*) la media, *b*) la varianza de la suma de los puntos.

3.108. Sea X una variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} cx & 0 \leq x \leq 2 \\ 0 & \text{si no es así} \end{cases}$$

donde c es una constante adecuada. Encuentre *a*) la media, *b*) la varianza, *c*) la función generadora de momentos, *d*) la función característica, *e*) el coeficiente de sesgo, *f*) el coeficiente de curtosis.

3.109. Si X y Y tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} cxy & 0 < x < 1, 0 < y < 1 \\ 0 & \text{si no es así} \end{cases}$$

Encuentre *a*) $E(X^2 + Y^2)$, *b*) $E(\sqrt{X^2 + Y^2})$.

3.110. Repita el problema 3.109 si X y Y son variables aleatorias independientes distribuidas de forma idéntica cuya función de densidad es $f(u) = (2\pi)^{-1/2} e^{-u^2/2}$, $-\infty < u < \infty$.

3.111. Sea X una variable aleatoria cuya función de densidad es

$$f(x) = \begin{cases} \frac{1}{2} & -1 < x < 1 \\ 0 & \text{si no es así} \end{cases}$$

y sea $Y = X^2$. Determine *a*) $E(X)$, *b*) $E(Y)$, *c*) $E(XY)$.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

/ / /

/

/ / / /

/ / / /

/ / / / / /

/ / / /

/ $\sqrt{35/12}$

/ $\sqrt{4/3}$

$\sigma_X = \sqrt{5}$ / $\sigma_X = \sqrt{15}/20$

/ / $\sqrt{15}/2$

$\frac{1}{2}(e^{t/2} + e^{-t/2}) = \cosh(t/2)$ $\mu = 0, \mu_2' = 1, \mu_3' = 0, \mu_4' = 1$

$+ 2te^{2t} - e^{2t})/2t^2$ $\mu = 4/3, \mu_2' = 2, \mu_3' = 16/5, \mu_4' = 16/3$

$\mu = 0, \mu_2 = 5, \mu_3 = -5, \mu_4 = 35$ $\mu = 0, \mu_2 = 3/80, \mu_3 = -121/160, \mu_4 = 2\,307/8\,960$

a) $1/(1-t), |t| < 1$ $\mu = 1, \mu_2' = 2, \mu_3' = 6, \mu_4' = 24$

$\mu_1 = 0, \mu_2 = 1, \mu_3 = 2, \mu_4 = 33$

$^{+1} - a^{k+1})/(k+1)(b-a)$ $+ (-1)^k](b-a)^k/2^{k+1}(k+1)$

$pe^{i\omega a} + qe^{i\omega b}$ **3.71.** $(\operatorname{sen} a\omega)/a\omega$ **3.72.** $(e^{2i\omega} - 2i\omega e^{2i\omega} - 1)/2\omega^2$

3.75. *a*) $11/144$ / $\sqrt{11}/12$ $\sqrt{11}/12$ / /

3.76.

3.77. / / $\sqrt{73/960}$ $\sqrt{73/960}$ / /

3.78. / / $\sqrt{233}/18$ $\sqrt{233}/18$ / /

3.79. $4/\sqrt{35}$ $-\sqrt{15}/4$

3.81. $+ 2)/(6x + 3)$ para $0 \le x \le 1$ $+ 2)/(6y + 3)$ para $0 \le y \le 1$

3.82. *a*) $1/2$ para $x \ge 0$ $\ge$

3.83.

X	0	1	2
$E(Y \mid X)$	$4/3$	1	$5/7$

Y	0	1	2
$E(X \mid Y)$	$4/3$	$7/6$	$1/2$

3.84. *a*) $\dfrac{6x^2 + 6x + 1}{18(2x + 1)^2}$ para $0 \le x \le 1$ $\dfrac{6y^2 + 6y + 1}{18(2y + 1)^2}$ $\le$ $\le$

3.85. /

3.86.

X	0	1	2
$\mathrm{Var}(Y \mid X)$	$5/9$	$4/5$	$24/49$

Y	0	1	2
$\mathrm{Var}(X \mid Y)$	$5/9$	$29/36$	$7/12$

3.87. *a*) $1/2$

3.92. $+ 0$ $1/\sqrt{3}$ $\sqrt{1 - (1/\sqrt{2})}$ /

3.94.

3.96. $1 - \frac{1}{2}\sqrt{3}$ /

3.97. $\sqrt{1 - (3/\sqrt{10})}$ $\sqrt{1 - (\sqrt{3}/2)}$ $\sqrt{1/2}$ $\sqrt{1 - (1/\sqrt{10})}$

3.98. $(\sqrt{3} - 1)/4$ /

3.99. –

3.101. $(5 - 2\sqrt{3})/3$ $(3 - 2e^{-1}\sqrt{3})/3$

3.102. /

3.105. *a*) $7/3$ / $+ 2e^{2t} + 3e^{3t})/6$ *d*) $(e^{i\omega} + 2e^{2i\omega} + 3e^{3i\omega})/6$ *e*) $-7/27$

3.106. *a*) $1/3$ / $-$ $-$ / $-2(e^{i\omega} - 1 - i\omega)/\omega^2$ /

3.107. / /

3.108. / / $+ 2te^{2t} - e^{2t})/2t^2$ $-(1 + 2i\omega e^{2i\omega} - e^{2i\omega})/2\omega^2$
$-2\sqrt{18}/15$ /

3.109. $8(2\sqrt{2} - 1)/15$

3.110. $\sqrt{2\pi}/2$

3.111. /

CAPÍTULO 4

Distribuciones especiales de probabilidad

LA DISTRIBUCIÓN BINOMIAL

Suponga que se realiza un experimento como el lanzamiento de una moneda o de un dado repetidas veces o la extracción, también repetidamente, de una canica de una urna. Cada lanzamiento o cada extracción se llama *ensayo*. En cada uno de ellos habrá una probabilidad de ocurrencia asociada a un evento particular, como, por ejemplo, obtener cara en el caso de la moneda, 4 en el caso del dado o extraer una canica roja. En algunos casos esta probabilidad no varía de un ensayo a otro (como en el caso del lanzamiento de la moneda o el dado). Tales ensayos son *independientes* y se les llama *ensayos de Bernoulli* en honor a James Bernoulli, quien los investigó a fines del siglo XVII.

Sea p la probabilidad de que ocurra un evento en un ensayo de Bernoulli (a lo que se le llama *probabilidad de éxito*). Entonces $q = 1 - p$ es la probabilidad de que en un ensayo ese evento no ocurra (a lo que se llama *probabilidad de fracaso*). La probabilidad de que el evento ocurra exactamente x veces en n ensayos (es decir, x éxitos y $n - x$ fracasos) está dada por la función de probabilidad

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x} \tag{1}$$

donde la variable aleatoria X denota el número de éxitos en n ensayos y $x = 0, 1, \ldots, n$.

EJEMPLO 4.1 La probabilidad de obtener exactamente 2 caras en 6 lanzamientos de una moneda legal (no cargada) es

$$P(X = 2) = \binom{6}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{6-2} = \frac{6!}{2!4!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{6-2} = \frac{15}{64}$$

A la función de probabilidad discreta (1) se le llama *distribución binomial* debido a que $x = 0, 1, 2, \ldots, n$, corresponde a los términos sucesivos de la *expansión binomial*

$$(q + p)^n = q^n + \binom{n}{1} q^{n-1}p + \binom{n}{2} q^{n-2}p^2 + \cdots + p^n = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} \tag{2}$$

Al caso especial de una distribución binomial en la que $n = 1$ se le conoce como *distribución de Bernoulli.*

ALGUNAS PROPIEDADES DE LA DISTRIBUCIÓN BINOMIAL

En la tabla 4-1 se listan algunas propiedades de la distribución binomial.

Tabla 4-1

Media	$\mu = np$
Varianza	$\sigma^2 = npq$
Desviación estándar	$\sigma = \sqrt{npq}$
Coeficiente de sesgo	$\alpha_3 = \dfrac{q - p}{\sqrt{npq}}$
Coeficiente de curtosis	$\alpha_4 = 3 + \dfrac{1 - 6pq}{npq}$
Función generadora de momentos	$M(t) = (q + pe^t)^n$
Función característica	$\phi(\omega) = (q + pe^{i\omega})^n$

EJEMPLO 4.2 En 100 lanzamientos de una moneda legal, la esperanza o media del número de caras es $\mu = (100)\left(\frac{1}{2}\right) = 50$, y la desviación estándar es $\sigma = \sqrt{(100)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 5$.

LEY DE LOS GRANDES NÚMEROS PARA ENSAYOS DE BERNOULLI

La ley de los grandes números, de la página 83, tiene una interpretación interesante en el caso de los ensayos de Bernoulli, que se presenta en el siguiente teorema.

Teorema 4-1 **(Ley de los grandes números para ensayos de Bernoulli)** Sea X una variable aleatoria que indica el número de éxitos en n ensayos de Bernoulli, de manera que X/n es la proporción de éxitos. Entonces, si p es la probabilidad de éxito y ϵ es cualquier número positivo,

$$\lim_{n\to\infty} P\left(\left|\frac{X}{n} - p\right| \geq \epsilon\right) = 0 \tag{3}$$

En otras palabras, a la larga es muy probable que la proporción de éxitos, X/n, esté tan cerca como se desee de la probabilidad de éxito, p, en un ensayo. Esta ley justifica, de alguna manera, el uso de la definición empírica de probabilidad dada en la página 5. Un resultado más consistente es el que proporciona la *ley fuerte* de los grandes números (página 83), que establece que con probabilidad uno, $\lim_{n\to\infty} X/n = p$, es decir, *$X/n$ converge a p* salvo en algún número pequeño de casos.

DISTRIBUCIÓN NORMAL

Uno de los ejemplos más importantes de distribuciones de probabilidad continua es la *distribución normal*, también llamada *distribución gaussiana.* La función de densidad de esta distribución es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty \tag{4}$$

donde μ y σ son la media y desviación estándar, respectivamente. La correspondiente función de distribución es

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(v-\mu)^2/2\sigma^2}\, dv \tag{5}$$

Si X tiene la función de distribución dada en (5), se considera que la variable aleatoria X está *normalmente distribuida* con media μ y varianza σ^2.

Si Z representa la variable estandarizada correspondiente a X, es decir si

$$Z = \frac{X - \mu}{\sigma} \tag{6}$$

entonces la media o valor esperado de Z es 0 y la varianza es 1. En este caso, la función de densidad para Z se obtiene de (4) sustituyendo formalmente $\mu = 0$ y $\sigma = 1$, con lo que se obtiene

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \tag{7}$$

Esta expresión suele conocerse como *función de densidad normal estándar.* La función de distribución correspondiente es

$$F(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2}\, du = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_{0}^{z} e^{-u^2/2}\, du \tag{8}$$

Algunas veces al valor z de la variable estandarizada Z se le llama *puntuación estándar.* La función $F(z)$ está relacionada con la *función del error*, $\text{erf}(z)$. Se tiene

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_{0}^{z} e^{-u^2}\, du \qquad \text{y} \qquad F(z) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{z}{\sqrt{2}}\right)\right] \tag{9}$$

En la figura 4-1 se muestra una gráfica de la función de densidad (7), que suele llamarse *curva normal estándar.* En esta gráfica se han indicado las áreas que no se encuentran a más de 1, 2 y 3 desviaciones estándar de la media (es decir, entre $z = -1$ y $z = +1$, $z = -2$ y $z = +2$ y $z = -3$ y $z = +3$), las cuales son, respectivamente, iguales a 68.27%, 95.45% y 99.73% del área total, que es uno. Esto significa que

$$P(-1 \le Z \le 1) = 0.6827, \qquad P(-2 \le Z \le 2) = 0.9545, \qquad P(-3 \le Z \le 3) = 0.9973 \tag{10}$$

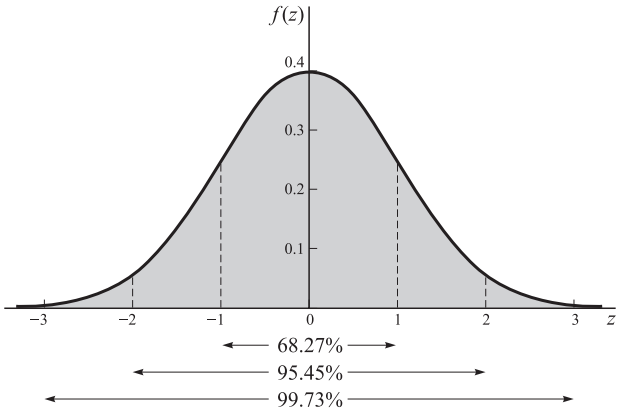


Figura 4-1

En el apéndice C se presenta una tabla que da las áreas bajo esta curva que están limitadas por las ordenadas $z = 0$ y cualquier valor positivo de z. Con esta tabla pueden encontrarse las áreas entre cualesquiera de las dos ordenadas aprovechando la simetría de la curva respecto a $z = 0$.

ALGUNAS PROPIEDADES DE LA DISTRIBUCIÓN NORMAL

En la tabla 4-2 se listan algunas propiedades importantes de la distribución normal.

Tabla 4-2

Media	μ
Varianza	σ^2
Desviación estándar	σ
Coeficiente de sesgo	$\alpha_3 = 0$
Coeficiente de curtosis	$\alpha_4 = 3$
Función generadora de momentos	$M(t) = e^{\mu t + (\sigma^2 t^2/2)}$
Función característica	$\phi(\omega) = e^{i\mu\omega - (\sigma^2\omega^2/2)}$

RELACIÓN ENTRE LAS DISTRIBUCIONES BINOMIAL Y NORMAL

Si n es grande, y si ni p ni q están demasiado cerca de cero, la distribución binomial puede aproximarse mediante una distribución normal cuya variable aleatoria estandarizada está dada por

$$Z = \frac{X - np}{\sqrt{npq}} \tag{11}$$

En este caso, X es la variable aleatoria que indica el número de éxitos en n ensayos de Bernoulli y p es la probabilidad de éxito. A medida que n aumenta, la aproximación mejora hasta ser exacta en el caso límite. (Vea el problema 4.17.) En la práctica, esta aproximación es muy buena si tanto np como nq son mayores a 5. El hecho de que la distribución binomial se aproxime a la distribución normal puede describirse de la siguiente manera

$$\lim_{n\to\infty} P\left(a \le \frac{X - np}{\sqrt{npq}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2}\, du \tag{12}$$

Es decir, la variable aleatoria estandarizada $(X - np)/\sqrt{npq}$ es *asintóticamente normal estándar.*

DISTRIBUCIÓN DE POISSON

Sea X una variable aleatoria discreta que puede tomar los valores 0, 1, 2 , . . . , de manera que la función de probabilidad de X está dada por

$$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, 2, \ldots \tag{13}$$

donde λ es una constante positiva. Esta distribución se llama *distribución de Poisson* (en honor a S. D. Poisson, quien la descubrió en la primera mitad del siglo XIX); una variable aleatoria que tiene esta distribución se considera que está distribuida de acuerdo con la *distribución de Poisson.*

Los valores de $f(x)$ en (13) pueden obtenerse usando el apéndice G, que da los valores de $e^{-\lambda}$ para varios valores de λ.

ALGUNAS PROPIEDADES DE LA DISTRIBUCIÓN DE POISSON

En la tabla 4-3 se listan algunas propiedades importantes de la distribución de Poisson.

Tabla 4-3

Mediana	$\mu = \lambda$
Varianza	$\sigma^2 = \lambda$
Desviación estándar	$\sigma = \sqrt{\lambda}$
Coeficiente de sesgo	$\alpha_3 = 1/\sqrt{\lambda}$
Coeficiente de curtosis	$\alpha_4 = 3 + (1/\lambda)$
Función generadora de momentos	$M(t) = e^{\lambda(e^t - 1)}$
Función característica	$\phi(\omega) = e^{\lambda(e^{i\omega} - 1)}$

RELACIÓN ENTRE LAS DISTRIBUCIONES BINOMIAL Y DE POISSON

En la distribución binomial (1), cuando n es grande y la probabilidad p de que ocurra un evento es cercana a cero, de manera que $q = 1 - p$ es cercana a 1, el evento recibe el nombre de *evento raro*. En la práctica se considera un evento raro si el número de ensayos es por lo menos 50 ($n \ge 50$) y np es menor a 5. En tales casos la distribución binomial puede aproximarse muy bien mediante la distribución de Poisson (13) con $\lambda = np$. Esto es de esperarse si se comparan las tablas 4-1 y 4-3, ya que sustituyendo en la tabla 4-1 $\lambda = np$, $q \approx 1$ y $p \approx 0$, se obtienen los resultados de la tabla 4-3.

RELACIÓN ENTRE LAS DISTRIBUCIONES DE POISSON Y LA NORMAL

Como existe una relación entre la distribución binomial y la normal, y también entre la binomial y la de Poisson, es de esperarse que exista también una relación entre la distribución de Poisson y la normal. En efecto, éste es el caso. Puede demostrarse que si X es la variable aleatoria de Poisson de (13) y $(X - \lambda)/\sqrt{\lambda}$ es la correspondiente variable aleatoria estandarizada, entonces

$$\lim_{\lambda\to\infty} P\left(a \le \frac{X-\lambda}{\sqrt{\lambda}} \le b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-u^2/2}\,du \tag{14}$$

es decir, la distribución de Poisson se aproxima a la distribución normal cuando $\lambda \to \infty$ o $(X - \lambda)/\sqrt{\lambda}$ es *asintóticamente normal estándar*.

TEOREMA DEL LÍMITE CENTRAL

La similitud entre (12) y (14) obliga a preguntarse si existen otras distribuciones además de la binomial y la de Poisson que tengan a la distribución normal como caso límite. El siguiente teorema indica que, en realidad, existe una numerosa clase de distribuciones que tienen esta propiedad.

Teorema 4-2 **(Teorema del límite central)** Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes distribuidas de manera idéntica (es decir, todas con la *misma* función de probabilidad en el caso discreto, o función de densidad, en el caso continuo) con media μ y varianza σ^2 finitas. Entonces, si $S_n = X_1 + X_2 + \cdots + X_n$ $(n = 1, 2 \ldots)$,

$$\lim_{n\to\infty} P\left(a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} \le b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-u^2/2}\,du \tag{15}$$

es decir, la variable aleatoria $(S_n - n\mu)/\sigma\sqrt{n}$, que es la variable estandarizada correspondiente a S_n, es asintóticamente normal estándar.

Este teorema también es válido bajo condiciones más generales; por ejemplo, cuando $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes con una misma media y una misma varianza, pero no necesariamente idénticamente distribuidas.

DISTRIBUCIÓN MULTINOMIAL

Suponga que los eventos $A_1, A_2, \ldots, A_k$ son mutuamente excluyentes y que pueden ocurrir con probabilidades respectivas $p_1, p_2, \ldots, p_k$, donde $p_1 + p_2 + \cdots + p_k = 1$. Si $X_1, X_2, \ldots, X_k$ son las variables aleatorias que dan el número de veces que ocurren $A_1, A_2, \ldots, A_k$ en un total de n ensayos, de manera que $X_1 + X_2 + \cdots + X_k = n$, entonces

$$P(X_1 = n_1, X_2 = n_2, \ldots, X_k = n_k) = \frac{n}{n_1!n_2!\cdots n_k!}\,p_1^{n_1}p_2^{n_2}\cdots p_k^{n_k} \tag{16}$$

donde $n_1 + n_2 + \cdots + n_k = n$, es la función de probabilidad conjunta de las variables aleatorias $X_1, \ldots, X_k$.

Esta distribución, que es una generalización de la distribución binomial, se llama *distribución multinomial*, ya que (16) es el término general de la expansión multinomial de $(p_1 + p_2 + \cdots + p_k)^n$.

EJEMPLO 4.3 Si un dado no cargado se lanza 12 veces, la probabilidad de obtener 1, 2, 3, 4, 5 y 6 exactamente dos veces cada uno es

$$P(X_1 = 2, X_2 = 2, \ldots, X_6 = 2) = \frac{12!}{2!2!2!2!2!2!}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2 = \frac{1\,925}{559\,872} = 0.00344$$

El número de veces que se espera ocurran $A_1, A_2, \ldots, A_k$ en n ensayos es $np_1, np_2, \ldots, np_k$, respectivamente, es decir,

$$E(X_1) = np_1, \qquad E(X_2) = np_2, \qquad \ldots, \qquad E(X_k) = np_k \tag{17}$$

DISTRIBUCIÓN HIPERGEOMÉTRICA

Suponga que una caja contiene b canicas azules y r rojas. Se realizan n ensayos de un experimento en el que en forma aleatoria se extrae una canica, se observa su color y se devuelve a la caja. A este tipo de experimentos suele

llamársele *muestreo con reemplazo.* En este caso, si X es la variable aleatoria que denota el número de canicas azules extraídas (éxitos) en n ensayos, entonces, usando la distribución binomial (1), se puede observar que la probabilidad de tener exactamente x éxitos es

$$P(X = x) = \binom{n}{x} \frac{b^x r^{n-x}}{(b + r)^n}, \qquad x = 0, 1, \ldots, n \tag{18}$$

ya que $p = b/(b + r)$, $q = 1 - p = r/(b + r)$.

Si lo anterior se modifica de manera que el *muestreo se haga sin reemplazo,* es decir, que las canicas no se devuelvan a la caja después de haber sido extraídas, entonces

$$P(X = x) = \frac{\binom{b}{x}\binom{r}{n-x}}{\binom{b+r}{n}}, \qquad x = \begin{matrix}\text{máx } (0, n - r), \ldots, \\ \text{mín } (n, b)\end{matrix} \tag{19}$$

Ésta es la *distribución hipergeométrica.* La media y la varianza de esta distribución son

$$\mu = \frac{nb}{b + r}, \qquad \sigma^2 = \frac{nbr(b + r - n)}{(b + r)^2 (b + r - 1)} \tag{20}$$

Si N denota el número total de canicas azules y rojas, mientras que las proporciones entre éstas se denotan por p y $q = 1 - p$, respectivamente, entonces

$$p = \frac{b}{b + r} = \frac{b}{N}, \qquad q = \frac{r}{b + r} = \frac{r}{N} \quad \text{o bien} \quad b = Np, \qquad r = Nq \tag{21}$$

de manera que (19) y (20) se convierten, respectivamente, en

$$P(X = x) = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}} \tag{22}$$

$$\mu = np, \qquad \sigma^2 = \frac{npq(N - n)}{N - 1} \tag{23}$$

Observe que cuando $N \to \infty$ (o bien N es grande en comparación con n), (22) se reduce a (18), lo que puede escribirse como

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \tag{24}$$

y (23) se reduce a

$$\mu = np, \qquad \sigma^2 = npq \tag{25}$$

en concordancia con los dos primeros renglones de la tabla 4-1, página 109. Los resultados son los que podían esperarse, ya que con N grande, el muestreo sin reemplazo es prácticamente idéntico al muestreo con reemplazo.

DISTRIBUCIÓN UNIFORME

Se dice que una variable aleatoria X está distribuida uniformemente en $a \le x \le b$ si su función de densidad es

$$f(x) = \begin{cases} 1/(b - a) & a \le x \le b \\ 0 & \text{de lo contrario} \end{cases} \tag{26}$$

distribución a la que se le llama *distribución uniforme*.

La función de distribución es

$$F(x) = P(X \le x) = \begin{cases} 0 & x < a \\ (x - a)/(b - a) & a \le x < b \\ 1 & x \ge b \end{cases} \tag{27}$$

La media y la varianza son, respectivamente,

$$\mu = \frac{1}{2}(a + b), \qquad \sigma^2 = \frac{1}{12}(b - a)^2 \tag{28}$$

DISTRIBUCIÓN DE CAUCHY

Una variable aleatoria X es una *distribución de Cauchy*, o tiene la *distribución de Cauchy* si la función de densidad de X es

$$f(x) = \frac{a}{\pi(x^2 + a^2)} \qquad a > 0, -\infty < x < \infty \tag{29}$$

Esta función de densidad es simétrica respecto de $x = 0$, de manera que la mediana es cero. Sin embargo, la media, la varianza y los momentos de orden superior no existen, lo cual también sucede con la función generadora de momentos; por otra parte, la que sí existe es la función característica, dada por

$$\phi(\omega) = e^{-a\omega} \tag{30}$$

DISTRIBUCIÓN GAMMA

Una variable aleatoria X tiene una *distribución gamma*, o es una *distribución gamma*, si su función de densidad se representa así:

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)} & x > 0 \\ 0 & x \le 0 \end{cases} \qquad (\alpha, \beta > 0) \tag{31}$$

donde $\Gamma(\alpha)$ es la *función gamma* (vea el apéndice A). La media y la varianza están dadas por

$$\mu = \alpha\beta, \qquad \sigma^2 = \alpha\beta^2 \tag{32}$$

La función generadora de momentos y la función característica están dadas, respectivamente, por

$$M(t) = (1 - \beta t)^{-\alpha}, \qquad \phi(\omega) = (1 - \beta i\omega)^{-\alpha} \tag{33}$$

DISTRIBUCIÓN BETA

Una variable aleatoria se llama *distribución beta*, o es una *distribución beta*, si la función de densidad tiene la forma

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}(1 - x)^{\beta-1}}{B(\alpha, \beta)} & 0 < x < 1 \\ 0 & \text{si no es así} \end{cases} \qquad (\alpha, \beta > 0) \tag{34}$$

donde $B(\alpha, \beta)$ es la *función beta* (vea el apéndice A). En vista de la relación (9), apéndice A, entre las funciones beta y gamma, la distribución beta también puede definirse mediante la función de densidad

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{si no es así} \end{cases} \tag{35}$$

donde α y β son positivas. La media y la varianza son

$$\mu = \frac{\alpha}{\alpha + \beta}, \qquad \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{36}$$

Para $\alpha > 1$, $\beta > 1$ existe una moda única de valor

$$x_{\text{moda}} = \frac{\alpha - 1}{\alpha + \beta - 2} \tag{37}$$

DISTRIBUCIÓN JI CUADRADA

Sean $X_1, X_2, \ldots, X_\nu$, ν variables aleatorias independientes distribuidas normalmente con media cero y varianza 1. Considere la variable aleatoria

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_\nu^2 \tag{38}$$

χ^2 se llama *ji cuadrada.* Puede demostrarse que para $x \geq 0$,

$$P(\chi^2 \leq x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^x u^{(\nu/2)-1} e^{-u/2}\, du \tag{39}$$

y $P(\chi^2 \leq x) = 0$ para $x < 0$.

A la función definida en (39) se le llama *distribución ji cuadrada,* y a ν *número de grados de libertad.* La distribución definida en (39) tiene como función de densidad la dada por

$$f(x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)}\, x^{(\nu/2)-1} e^{-x/2} & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{40}$$

Se puede observar que la distribución ji cuadrada es un caso especial de la distribución gamma, en la que $\alpha = \nu/2$, $\beta = 2$. Por tanto,

$$\mu = \nu, \qquad \sigma^2 = 2\nu, \qquad M(t) = (1 - 2t)^{-\nu/2}, \qquad \phi(\omega) = (1 - 2i\omega)^{-\nu/2} \tag{41}$$

Para ν grande ($\nu \geq 30$), se puede demostrar que $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ tiene una distribución casi normal con media 0 y varianza 1.

Tres teoremas que serán útiles más adelante son los siguientes:

Teorema 4-3 Sean $X_1, X_2, \ldots, X_\nu$ variables aleatorias independientes distribuidas normalmente con media 0 y varianza 1. Entonces, $\chi^2 = X_1^2 + X_2^2 + \cdots + X_\nu^2$ tiene la distribución ji cuadrada con ν grados de libertad.

Teorema 4-4 Sean $U_1, U_2, \ldots, U_k$ variables aleatorias independientes que tienen la distribución ji cuadrada con $\nu_1, \nu_2, \ldots, \nu_k$ grados de libertad, respectivamente. Entonces su suma $W = U_1 + U_2 + \cdots + U_k$ tiene la distribución ji cuadrada con $\nu_1 + \nu_2 + \cdots + \nu_k$ grados de libertad.

Teorema 4-5 Sean V_1 y V_2 variables aleatorias independientes. Suponga que V_1 tiene la distribución ji cuadrada con ν_1 grados de libertad, mientras que $V = V_1 + V_2$ tiene la distribución ji cuadrada con ν grados de libertad, donde $\nu > \nu_1$. Entonces, V_2 tiene la distribución ji cuadrada con $\nu - \nu_1$ grados de libertad.

En los trabajos de estadística, cuando se trata de la distribución ji cuadrada, la distribución *t* (a continuación), la distribución *F* (página 116) y otras, es común que se use el *mismo símbolo* tanto para la variable aleatoria como para un valor de ella. Por tanto, los valores percentiles de la distribución ji cuadrada con ν grados de libertad se denotan con $\chi^2_{p,\nu}$ o simplemente χ^2_p, cuando ν se sobrentiende, y no $x_{p,\nu}$ o x_p. (Vea el apéndice E.) Ésta es una notación ambigua y el lector deberá tener cuidado con ella, especialmente cuando cambian las variables en la función de densidad.

DISTRIBUCIÓN *t* DE STUDENT

Si una variable aleatoria tiene la función de densidad

$$f(t) = \frac{\Gamma\left(\dfrac{\nu + 1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\dfrac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2} \qquad -\infty < t < \infty \tag{42}$$

se llama *distribución t de Student*, con ν grados de libertad. Si ν es grande ($\nu \geq 30$), la gráfica de $f(t)$ se aproxima a la curva normal estándar, como se indica en la figura 4-2. Los valores percentiles de la distribución *t* para ν grados

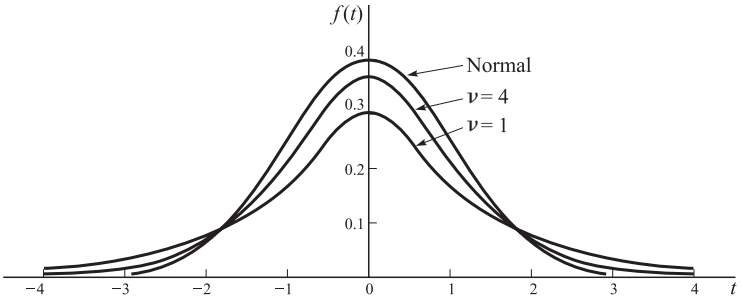


Figura 4-2

de libertad se denotan por $t_{p,\nu}$, o simplemente t_p cuando ν se sobrentiende. En el apéndice A se presenta una tabla que contiene estos valores. Como la distribución t es simétrica, $t_{1-p} = -t_p$; por ejemplo $t_{0.05} = -t_{0.95}$.

En la distribución t se tiene

$$\mu = 0 \qquad \text{y} \qquad \sigma^2 = \frac{\nu}{\nu - 2} \qquad (\nu > 2). \tag{43}$$

El siguiente teorema será importante posteriormente.

Teorema 4-6 Sean Y y Z variables aleatorias independientes, donde Y está normalmente distribuida con media 0 y varianza 1, y Z sigue la distribución ji cuadrada con ν grados de libertad. Entonces, la variable aleatoria

$$T = \frac{Y}{\sqrt{Z/\nu}} \tag{44}$$

tiene distribución t con ν grados de libertad.

DISTRIBUCIÓN *F*

Una variable aleatoria tiene *distribución F* (nombre que se le da en honor a R. A. Fisher) *con* ν_1 *y* ν_2 *grados de libertad* si su función de densidad está dada por

$$f(u) = \begin{cases} \dfrac{\Gamma\left(\dfrac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)} \nu_1^{\nu_1/2}\nu_2^{\nu_2/2}u^{(\nu_1/2)-1}(\nu_2 + \nu_1 u)^{-(\nu_1+\nu_2)/2} & u > 0 \\ 0 & u \le 0 \end{cases} \tag{45}$$

Los valores percentiles de la distribución F para ν_1 y ν_2 grados de libertad se denotan F_{p,ν_1,ν_2}, o simplemente F_p cuando ν_1 y ν_2 se sobrentienden. En el apéndice F se presenta una tabla que da estos valores para $p = 0.95$ y $p = 0.99$.

La media y la varianza están dadas, respectivamente, por

$$\mu = \frac{\nu_2}{\nu_2 - 2} \ (\nu_2 > 2) \qquad \text{y} \qquad \sigma^2 = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)^2} \qquad (\nu_2 > 4) \tag{46}$$

Esta distribución tiene una sola moda en el valor

$$u_{\text{moda}} = \left(\frac{\nu_1 - 2}{\nu_1}\right)\left(\frac{\nu_2}{\nu_2 + 2}\right) \qquad (\nu_1 > 2) \tag{47}$$

Los siguientes teoremas serán importantes para lo que sigue.

Teorema 4-7 Sean V_1 y V_2 variables aleatorias independientes con distribución ji cuadrada con ν_1 y ν_2 grados de libertad, respectivamente. Entonces la variable aleatoria

$$V = \frac{V_1/\nu_1}{V_2/\nu_2} \tag{48}$$

tiene distribución F con ν_1 y ν_2 grados de libertad.

Teorema 4-8

$$F_{1-p,\nu_2,\nu_1} = \frac{1}{F_{p,\nu_1,\nu_2}}$$

RELACIÓN ENTRE LAS DISTRIBUCIONES JI CUADRADA, *t* Y *F*

Teorema 4-9

$$F_{1-p,1,\nu} = t^2_{1-(p/2),\nu}$$

Teorema 4-10

$$F_{p,\nu,\infty} = \frac{\chi^2_{p,\nu}}{\nu}$$

DISTRIBUCIÓN NORMAL BIVARIADA

Una generalización de la distribución normal para dos variables aleatorias continuas X y Y se da mediante la función de densidad conjunta

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right] \Big/ 2(1-\rho^2)\right\} \tag{49}$$

donde $-\infty < x < \infty$, $-\infty < y < \infty$; μ_1, μ_2 son las medias de X y Y; σ_1, σ_2 son las desviaciones estándar de X y de Y, mientras que ρ es el coeficiente de correlación entre X y Y. A (49) suele conocérsele como *distribución normal bivariada.*

En toda distribución conjunta, la condición $\rho = 0$ es necesaria para la independencia de las variables aleatorias (vea el teorema 3-15). En el caso de (49) esta condición también es suficiente (vea el problema 4.51).

DIVERSAS DISTRIBUCIONES

En las distribuciones que se enumeran a continuación, las constantes α, β, a, b, . . . se consideran positivas a menos que se indique lo contrario. La función característica $\phi(\omega)$ se obtiene de la función generadora de momentos, cuando ésta es dada, y se hace $t = i\omega$.

1. DISTRIBUCIÓN GEOMÉTRICA.

$$f(x) = P(X = x) = pq^{x-1} \qquad x = 1, 2, \ldots$$

$$\mu = \frac{1}{p} \qquad \sigma^2 = \frac{q}{p^2} \qquad M(t) = \frac{pe^t}{1 - qe^t}$$

La variable aleatoria X representa el número de ensayos de Bernoulli hasta el ensayo en el que se obtiene el primer éxito. En este caso, p es la probabilidad de éxito en un solo ensayo.

2. DISTRIBUCIÓN DE PASCAL O DISTRIBUCIÓN BINOMIAL NEGATIVA.

$$f(x) = P(X = x) = \binom{x-1}{r-1} p^r q^{x-r} \qquad x = r, r+1, \ldots$$

$$\mu = \frac{r}{p} \qquad \sigma^2 = \frac{rq}{p^2} \qquad M(t) = \left(\frac{pe^t}{1 - qe^t}\right)^r$$

La variable aleatoria X representa el número de ensayos de Bernoulli hasta que se obtiene el éxito r-ésimo, inclusive. El caso especial $r = 1$ proporciona la distribución geométrica.

3. DISTRIBUCIÓN EXPONENCIAL.

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

$$\mu = \frac{1}{\alpha} \qquad \sigma^2 = \frac{1}{\alpha^2} \qquad M(t) = \frac{\alpha}{\alpha - t}$$

4. DISTRIBUCIÓN DE WEIBULL.

$$f(x) = \begin{cases} abx^{b-1}e^{-ax^b} & x > 0 \\ 0 & x \le 0 \end{cases}$$

$$\mu = a^{-1/b}\Gamma\left(1 + \frac{1}{b}\right) \qquad \sigma^2 = a^{-2/b}\left[\Gamma\left(1 + \frac{2}{b}\right) - \Gamma^2\left(1 + \frac{1}{b}\right)\right]$$

5. DISTRIBUCIÓN DE MAXWELL.

$$f(x) = \begin{cases} \sqrt{2/\pi}\alpha^{3/2}x^2e^{-\alpha x^2/2} & x > 0 \\ 0 & x \le 0 \end{cases}$$

$$\mu = 2\sqrt{\frac{2}{\pi\alpha}} \qquad \sigma^2 = \left(3 - \frac{8}{\pi}\right)\alpha^{-1}$$

PROBLEMAS RESUELTOS

DISTRIBUCIÓN BINOMINAL

4.1. Determinar la probabilidad de que en tres lanzamientos de una moneda legal se obtenga *a*) 3 caras, *b*) 2 cruces y 1 cara, *c*) por lo menos 1 cara, *d*) no más de 1 cruz.

Método 1

Sea *H* cara y *T* cruz, así que *HTH* significa cara en el primer lanzamiento, cruz en el segundo y cara en el tercero.

Como en cada lanzamiento hay dos posibilidades (cara o cruz) en total hay (2)(2)(2) = 8 posibles resultados, es decir, puntos muestrales en el espacio muestral. Éstos son

$$HHH, HHT, HTH, HTT, TTH, THH, THT, TTT$$

En el caso de una moneda legal, a estos resultados se les asigna probabilidades iguales de $1/8$. Por tanto,

a) $P(\text{3 caras}) = P(HHH) = \frac{1}{8}$

b) $P(\text{2 cruces y 1 cara}) = P(HTT \cup TTH \cup THT)$

$$= P(HTT) + P(TTH) + P(THT) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

c) $P(\text{por lo menos 1 cara})$

$= P(1, 2 \text{ o } 3 \text{ caras})$

$= P(\text{1 cara}) + P(\text{2 caras}) + P(\text{3 caras})$

$= P(HTT \cup THT \cup TTH) + P(HHT \cup HTH \cup THH) + P(HHH)$

$= P(HTT) + P(THT) + P(TTH) + P(HHT) + P(HTH) + P(THH) + P(HHH) = \frac{7}{8}$

De manera alternativa,

$$P(\text{al menos 1 cara}) = 1 - P(\text{ninguna cara}) = 1 - P(TTT) = 1 - \frac{1}{8} = \frac{7}{8}$$

d) $P(\text{no más de 1 cruz}) = P(\text{0 cruces o 1 cruz})$

$= P(\text{0 cruz}) + P(\text{1 cruz})$

$= P(HHH) + P(HHT \cup HTH \cup THH)$

$= P(HHH) + P(HHT) + P(HTH) + P(THH)$

$= \frac{4}{8} = \frac{1}{2}$

Método 2 (usando fórmula)

a) $P(3 \text{ caras}) = \binom{3}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 = \frac{1}{8}$

b) $P(2 \text{ cruces y } 1 \text{ cara}) = \binom{3}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1 = \frac{3}{8}$

c) $P(\text{por lo menos 1 cara}) = P(1, 2, \text{o } 3 \text{ caras})$

$= P(1 \text{ cara}) + P(2 \text{ caras}) + P(3 \text{ caras})$

$= \binom{3}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2 + \binom{3}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1 + \binom{3}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 = \frac{7}{8}$

De manera alternativa,

$P(\text{por lo menos 1 cara}) = 1 - P(\text{ninguna cara})$

$= 1 - \binom{3}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^3 = \frac{7}{8}$

d) $P(\text{no más de 1 cruz}) = P(0 \text{ cruces o } 1 \text{ cruz})$

$= P(0 \text{ cruces}) + P(1 \text{ cruz})$

$= \binom{3}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 + \binom{3}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right) = \frac{1}{2}$

Cabe indicar que también podemos usar la notación de variables aleatorias. Por ejemplo, si X es la variable aleatoria que denota la cantidad de caras en 3 lanzamientos, *c*) lo escribimos como

$$P(\text{por lo menos 1 cara}) = P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{7}{8}$$

En este caso, estas dos formas se usarán de manera indistinta.

4.2. Determinar la probabilidad de que en cinco lanzamientos de un dado no cargado se obtenga un 3 *a*) dos veces, *b*) a lo sumo una vez, *c*) por lo menos dos veces.

Sea X la variable aleatoria que indica la cantidad de veces que se obtiene 3 en cinco lanzamientos de un dado no cargado. Se tiene

$$\text{Probabilidad de obtener 3 en un solo lanzamiento} = p = \frac{1}{6}$$

$$\text{Probabilidad de no obtener 3 en un solo lanzamiento} = q = 1 - p = \frac{5}{6}$$

a) $P(\text{se obtenga 3 dos veces}) = P(X = 2) = \binom{5}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3 = \frac{625}{3\,888}$

b) $P(\text{se obtenga 3 a lo sumo una vez}) = P(X \leq 1) = P(X = 0) + P(X = 1)$

$$= \binom{5}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^5 + \binom{5}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^4$$

$$= \frac{3\,125}{7\,776} + \frac{3\,125}{7\,776} = \frac{3\,125}{3\,888}$$

c) $P(\text{se obtenga 3 por lo menos dos veces})$

$= P(X \geq 2)$

$= P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$= \binom{5}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^3 + \binom{5}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 + \binom{5}{4}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^1 + \binom{5}{5}\left(\frac{1}{6}\right)^5\left(\frac{5}{6}\right)^0$

$= \frac{625}{3\,888} + \frac{125}{3\,888} + \frac{25}{7\,776} + \frac{1}{7\,776} = \frac{763}{3\,888}$

4.3. Encontrar la probabilidad de que en una familia con 4 hijos haya *a*) por lo menos 1 niño, *b*) por lo menos 1 un niño y por lo menos 1 niña. La probabilidad de que nazca un varón es de $1/2$.

a) $P(1 \text{ niño}) = \binom{4}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3 = \frac{1}{4}, \qquad P(2 \text{ niños}) = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = \frac{3}{8}$

$P(3 \text{ niños}) = \binom{4}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1 = \frac{1}{4}, \qquad P(4 \text{ niños}) = \binom{4}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0 = \frac{1}{16}$

Entonces,

$$P(\text{por lo menos 1 niño}) = P(1 \text{ niño}) + P(2 \text{ niños}) + P(3 \text{ niños}) + P(4 \text{ niños})$$
$$= \frac{1}{4} + \frac{3}{8} + \frac{1}{4} + \frac{1}{16} = \frac{15}{16}$$

Otro método

$$P(\text{por lo menos 1 niño}) = 1 - P(\text{ningún niño}) = 1 - \left(\frac{1}{2}\right)^4 = 1 - \frac{1}{16} = \frac{15}{16}$$

b) $P(\text{por lo menos 1 niño y por lo menos 1 niña}) = 1 - P(\text{ningún niño}) - P(\text{ninguna niña})$

$$= 1 - \frac{1}{16} - \frac{1}{16} = \frac{7}{8}$$

Este problema también se puede resolver si se representa con X la variable aleatoria que denota el número de niños en una familia con cuatro hijos. Entonces, por ejemplo, *a*) será

$$P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = \frac{15}{16}$$

4.4. De 2 000 familias que tienen 4 hijos cada una, ¿cuántas se espera que tengan *a*) por lo menos 1 niño, *b*) exactamente 2 niños, *c*) 1 o 2 niñas, *d*) ninguna niña?

En el problema 4.3 se puede observar que

a) El número esperado de familias con por lo menos 1 niño $= 2\,000\left(\frac{15}{16}\right) = 1\,875$

b) El número esperado de familias con exactamente 2 niños $= 2\,000 \cdot P(2 \text{ niños}) = 2\,000\left(\frac{3}{8}\right) = 750$

c) $P(1 \text{ o } 2 \text{ niñas}) = P(1 \text{ niña}) + P(2 \text{ niñas})$

$$= P(1 \text{ niño}) + P(2 \text{ niños}) = \frac{1}{4} + \frac{3}{8} = \frac{5}{8}$$

El número esperado de familias que tengan una o dos niñas $= (2\,000)\left(\frac{5}{8}\right) = 1\,250$

d) El número esperado de familias en las que no haya ninguna niña $= (2\,000)\left(\frac{1}{16}\right) = 125$

4.5. Si 20% de los tornillos que produce una máquina están defectuosos, determinar la probabilidad que de 4 tornillos elegidos al azar, *a*) 1, *b*) 0, *c*) menos de 2, estén defectuosos.

La probabilidad de que un tornillo esté defectuoso es $p = 0.2$, y de que no lo sea es $q = 1 - p = 0.8$. Sea X la variable aleatoria que da el número de tornillos defectuosos. Entonces,

a) $P(X = 1) = \binom{4}{1}(0.2)^1(0.8)^3 = 0.4096$

b) $P(X = 0) = \binom{4}{0}(0.2)^0(0.8)^4 = 0.4096$

c) $P(X < 2) = P(X = 0) + P(X = 1)$

$$= 0.4096 + 0.4096 = 0.8192$$

4.6. Determinar la probabilidad de que en tres lanzamientos de un par de dados no cargados, por lo menos una vez el total sea 7.

En un solo lanzamiento de un par de dados no cargados la probabilidad de un 7 es $p = 1/6$ (vea el problema 2.1, página 44), de manera que la probabilidad de ningún 7 en un solo lanzamiento es $q = 1 - p = 5/6$. Entonces,

$$P(\text{ningún 7 en tres lanzamientos}) = \binom{3}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^3 = \frac{125}{216}$$

y
$$P(\text{por lo menos un 7 en tres lanzamientos}) = 1 - \frac{125}{216} = \frac{91}{216}$$

4.7. Determinar la función generadora de momentos de una variable aleatoria X con distribución binomial.

Método 1
Si X está distribuida de manera binomial,

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}$$

La función generadora de momentos está dada por

$$\begin{aligned} M(t) = E(e^{tx}) &= \sum e^{tx} f(x) \\ &= \sum_{x=0}^{n} e^{tx}\binom{n}{x} p^x q^{n-x} \\ &= \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x q^{n-x} \\ &= (q + pe^t)^n \end{aligned}$$

Método 2
Dada una secuencia de n ensayos de Bernoulli, se define

$$X_j = \begin{cases} 0 & \text{si el fracaso ocurre en el ensayo } j\text{-ésimo} \\ 1 & \text{si el éxito ocurre en el ensayo } j\text{-ésimo} \end{cases} \qquad (j = 1, 2, \ldots, n)$$

Entonces, las X_j son independientes y $X = X_1 + X_2 + \cdots + X_n$. Para la función generadora de momentos de X_j, tenemos

$$M_j(t) = e^{t0}q + e^{t1}p = q + pe^t \qquad (j = 1, 2, \ldots, n)$$

De acuerdo con el teorema 3-9 de la página 80,

$$M(t) = M_1(t)M_2(t)\ldots M_n(t) = (q + pe^t)^n$$

4.8. Demostrar que la media y la varianza de una variable aleatoria distribuida de manera binomial son, respectivamente, $\mu = np$ y $\sigma^2 = npq$.

Procediendo como en el método 2 del problema 4.7, para $j = 1, 2, \ldots, n$, se tiene

$$E(X_j) = 0q + 1p = p$$

$$\text{Var}(X_j) = E[(X_j - p)^2] = (0 - p)^2 q + (1 - p)^2 p$$

$$= p^2 q + q^2 p = pq(p + q) = pq$$

Por consiguiente,
$$\mu = E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = np$$

$$\sigma^2 = \text{Var}(X) = \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n) = npq$$

donde para σ^2 se ha usado el teorema 3-7.

Los resultados anteriores también pueden obtenerse (pero con más dificultad) diferenciando la función generadora de momento (vea el problema 3.38), o directamente de la función de probabilidad.

4.9. Si la probabilidad de que un tornillo esté defectuoso es 0.1, calcular *a*) la media y *b*) la desviación estándar del número de tornillos defectuosos de un total de 400 tornillos.

a) Media $\mu = np = (400)(0.1) = 40$, es decir, puede *esperarse* que 40 tornillos tengan algún defecto.
b) Varianza $\sigma^2 = npq = (400)(0.1)(0.9) = 36$. Por tanto, la desviación estándar es $\sigma = \sqrt{36} = 6$.

LEY DE LOS GRANDES NÚMEROS PARA ENSAYOS DE BERNOULLI

4.10. Demostrar el teorema 4-1, ley (débil) de los grandes números para ensayos de Bernoulli.

De acuerdo con la desigualdad de Chebyshev, página 83, si X es una variable aleatoria cualquiera, con media μ y varianza σ^2 finitas, entonces

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \tag{1}$$

En particular, si X es una distribución binomial o de Bernoulli, entonces $\mu = np$, $\sigma = \sqrt{npq}$ y (1) se transforma en

$$P(|X - np| \geq k\sqrt{npq}) \leq \frac{1}{k^2} \tag{2}$$

o bien

$$P\left(\left|\frac{X}{n} - p\right| \geq k\sqrt{\frac{pq}{n}}\right) \leq \frac{1}{k^2} \tag{3}$$

Si $\epsilon = k\sqrt{\dfrac{pq}{n}}$, (3) se transforma en

$$P\left(\left|\frac{X}{n} - p\right| \geq \epsilon\right) \leq \frac{pq}{n\epsilon^2}$$

y tomando el límite cuando $n \to \infty$ tenemos, como deseábamos,

$$\lim_{n\to\infty} P\left(\left|\frac{X}{n} - p\right| \geq \epsilon\right) = 0$$

Este resultado también es consecuencia directa del teorema 3-19, página 83, con $S_n = X$, $\mu = np$, $\sigma = \sqrt{npq}$.

4.11. Dar una interpretación de la ley (débil) de los grandes números para la ocurrencia de 3 lanzamientos sucesivos de un dado no cargado.

En este caso, la ley de los grandes números indica que la probabilidad de que la proporción de números 3 en n lanzamientos difiera de $1/6$ en más de cualquier valor $\epsilon > 0$ tiende a cero cuando $n \to \infty$.

DISTRIBUCIÓN NORMAL

4.12. Encontrar el área bajo la curva normal estándar que se muestra en la figura 4-3, *a*) entre $z = 0$ y $z = 1.2$, *b*) entre $z = -0.68$ y $z =$ cero, *c*) entre $z = -0.46$ y $z = 2.21$, *d*) entre $z = 0.81$ y $z = 1.94$, *e*) a la derecha de $z = -1.28$.

a) Con base en la tabla del apéndice C, bajamos por la columna marcada con z hasta la entrada 1.2. Después avanzamos hacia la derecha hasta la columna marcada 0. El resultado es 0.3849, que es el área que buscamos, representa la probabilidad de que Z esté entre 0 y 1.2 (figura 4-3). Por tanto,

$$P(0 \leq Z \leq 1.2) = \frac{1}{\sqrt{2\pi}}\int_0^{1.2} e^{-u^2/2}\,du = 0.3849$$

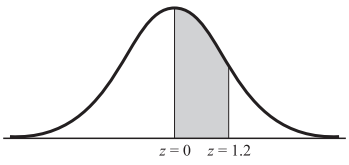


Figura 4-3

b) El área que buscamos está entre $z = 0$ y $z = +0.68$ (por simetría). Por tanto, en la columna marcada con z bajamos hasta llegar a la entrada 0.6. Después avanzamos hacia la derecha hasta la columna marcada con 8.

El resultado es 0.2517, que es el área que buscamos, y representa la probabilidad de que Z esté entre -0.68 y 0 (figura 4-4). Por tanto,

$$P(-0.68 \le Z \le 0) = \frac{1}{\sqrt{2\pi}}\int_{-0.68}^{0} e^{-u^2/2}\,du$$
$$= \frac{1}{\sqrt{2\pi}}\int_{0}^{0.68} e^{-u^2/2}\,du = 0.2517$$

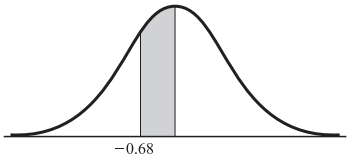


Figura 4-4

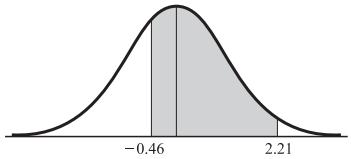


Figura 4-5

c) Área que buscamos = (área entre $z = -0.46$ y $z = 0$)
+ (área entre $z = 0$ y $z = 2.21$)
= (área entre $z = 0$ y $z = 0.46$)
+ (área entre $z = 0$ y $z = 2.21$)
$= 0.1772 + 0.4864 = 0.6636$

El área es 0.6636, y representa la probabilidad de que Z esté entre -0.46 y 2.21 (figura 4-5). Por tanto,

$$P(-0.46 \le Z \le 2.21) = \frac{1}{\sqrt{2\pi}}\int_{-0.46}^{2.21} e^{-u^2/2}\,du$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-0.46}^{0} e^{-u^2/2}\,du + \frac{1}{\sqrt{2\pi}}\int_{0}^{2.21} e^{-u^2/2}\,du$$
$$= \frac{1}{\sqrt{2\pi}}\int_{0}^{0.46} e^{-u^2/2}\,du + \frac{1}{\sqrt{2\pi}}\int_{0}^{2.21} e^{-u^2/2}\,du = 0.1772 + 0.4864$$
$$= 0.6636$$

d) El área que buscamos (figura 4-6) = (área entre $z = 0$ y $z = 1.94$)
− (área entre $z = 0$ y $z = 0.81$)
$= 0.4738 - 0.2910 = 0.1828$

Esto es lo mismo que $P(0.81 \le Z \le 1.94)$.

e) El área que buscamos (figura 4-7) = (área entre $z = -1.28$ y $z = 0$)
+ (área a la derecha de $z = 0$)
$= 0.3997 + 0.5 = 0.8997$

Esto es lo mismo que $P(Z \ge -1.28)$.

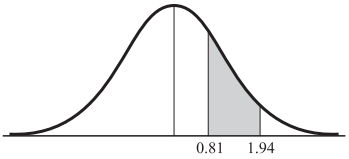


Figura 4-6

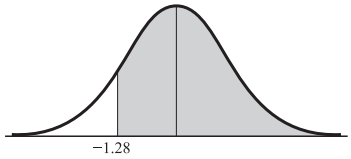


Figura 4-7

4.13. Si "área" se refiere a la zona bajo la curva normal estándar, calcular el valor o los valores de z tales que *a*) el área entre 0 y z es 0.3770, *b*) el área a la izquierda de z es 0.8621, *c*) el área entre -1.5 y z es 0.0217.

a) En la tabla del apéndice C, la entrada 0.3770 se localiza a la derecha del renglón marcado con 1.1 y debajo de la columna marcada con 6. Entonces, la z que buscamos es de 1.16.

Por simetría, -1.16 es otro valor de z. Por tanto, $z = \pm 1.16$ (figura 4-8). Este problema es equivalente a despejar z en la ecuación

$$\frac{1}{\sqrt{2\pi}}\int_0^z e^{-u^2/2}\,du = 0.3770$$

b) Como el área es mayor que 0.5, z debe ser positiva.

El área entre 0 y z es 0.8621 – 0.5 = 0.3621, de donde $z = 1.09$ (figura 4-9).

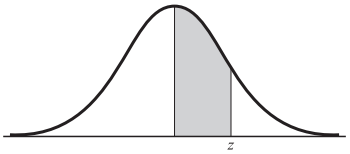


Figura 4-8

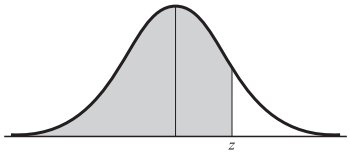


Figura 4-9

c) Si z fuera positiva, el área sería mayor que el área entre -1.5 y 0, que es 0.4332; de manera que z debe ser negativa.

Caso 1 z es negativa pero a la derecha de –1.5 (figura 4-10).

$$\begin{aligned}\text{Área entre } -1.5 \text{ y } z &= (\text{área entre } -1.5 \text{ y } 0)\\ &\quad - (\text{área entre } 0 \text{ y } z)\\ 0.0217 &= 0.4332 - (\text{área entre } 0 \text{ y } z)\end{aligned}$$

Entonces, el área entre 0 y z es 0.4332 – 0.0217 = 0.4115, de donde $z = -1.35$.

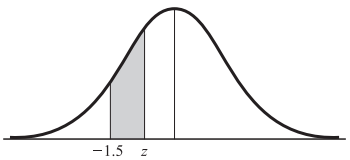


Figura 4-10

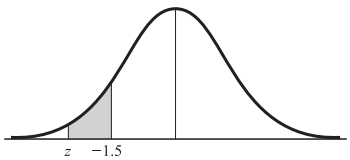


Figura 4-11

Caso 2 z es negativa pero a la izquierda de -1.5 (figura 4-11).

$$\begin{aligned}\text{Área entre } z \text{ y } -1.5 &= (\text{área entre } z \text{ y } 0)\\ &\quad - (\text{área entre } -1.5 \text{ y } 0)\\ 0.0217 &= (\text{área entre } 0 \text{ y } z) - 0.4332\end{aligned}$$

Entonces, el área entre 0 y z es $0.0217 + 0.4332 = 0.4549$ y $z = -1.694$ usando interpolación lineal; o, con un poco menos de precisión, $z = -1.69$.

4.14. El peso medio de 500 estudiantes varones de una universidad es de 151 lb y la desviación estándar es de 15 libras. Si se supone que el peso está distribuido normalmente, encontrar cuántos estudiantes pesan *a*) entre 120 y 155 libras, *b*) más de 185 libras.

a) Los pesos registrados entre 120 y 155 libras realmente pueden tener un valor desde 119.5 hasta 155.5 lb, suponiendo que estos pesos están dados a la libra más cercana (figura 4-12).

$$\begin{aligned}119.5 \text{ lb en unidades estándar} &= (119.5 - 151)/15\\ &= -2.10\\ 155.5 \text{ lb en unidades estándar} &= (155.5 - 151)/15\\ &= 0.30\end{aligned}$$

$$\begin{aligned}\text{Proporción de estudiantes que buscamos} &= (\text{área entre } z = -2.10 \text{ y } z = 0.30)\\ &= (\text{área entre } z = -2.10 \text{ y } z = 0)\\ &\quad + (\text{área entre } z = 0 \text{ y } z = 0.30)\\ &= 0.4821 + 0.1179 = 0.6000\end{aligned}$$

Por tanto, el número de estudiantes cuyo peso está entre 120 y 155 libras es 500(0.6000) = 300

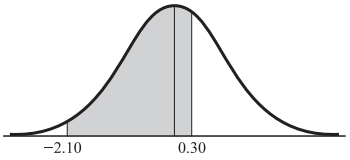


Figura 4-12

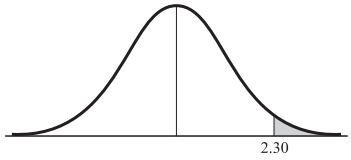


Figura 4-13

b) Los estudiantes que pesan más de 185 libras deben pesar por lo menos 185.5 libras (figura 4-13).

$$185.5 \text{ lb, en unidades estándar} = (185.5 - 151)/15 = 2.30$$

$$\begin{aligned}&\text{Proporción de estudiantes que buscamos}\\ &= (\text{área a la derecha de } z = 2.30)\\ &= (\text{área a la derecha de } z = 0)\\ &\quad - (\text{área entre } z = 0 \text{ y } z = 2.30)\\ &= 0.5 - 0.4893 = 0.0107\end{aligned}$$

Entonces el número de estudiantes cuyo peso es superior a 185 libras es 500(0.0107) = 5.

Si W denota el peso de un estudiante elegido al azar, los resultados anteriores pueden resumirse en términos de probabilidad escribiendo

$$P(119.5 \le W \le 155.5) = 0.6000 \qquad P(W \ge 185.5) = 0.0107$$

4.15. El diámetro interno medio de 200 rondanas fabricadas por una máquina es de 0.502 pulgadas y la desviación estándar es de 0.005 pulgadas. La máxima tolerancia para el diámetro de estas rondanas es de 0.496 a 0.508 pulgadas; de lo contrario, las rondanas se consideran defectuosas. Determinar el porcentaje de rondanas defectuosas producidas por esta máquina, suponiendo que los diámetros están distribuidos normalmente.

$$0.496 \text{ en unidades estándar} = (0.496 - 0.502)/0.005 = -1.2$$
$$0.508 \text{ en unidades estándar} = (0.508 - 0.502)/0.005 = 1.2$$

Proporción de rondanas no defectuosas

$$\begin{aligned}&= (\text{área bajo la curva normal entre } z = -1.2 \text{ y } z = 1.2)\\ &= (\text{doble del área entre } z = 0 \text{ y } z = 1.2)\\ &= 2(0.3849) = 0.7698 \text{ o } 77\%\end{aligned}$$

Por tanto, el porcentaje de rondanas defectuosas es 100% − 77% = 23% (figura 4-14).

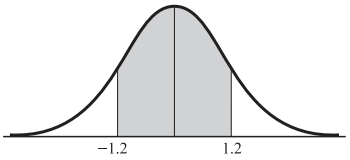


Figura 4-14

Obsérvese que si consideramos el intervalo de 0.496 a 0.508 pulgadas como representante de los diámetros que van de 0.4955 a 0.5085 pulgadas, el resultado anterior se modifica de manera casi imperceptible. Sin embargo, con dos cifras decimales, el resultado es el mismo.

4.16. Calcular la función generadora de momentos correspondiente a la distribución normal general.

Tenemos

$$M(t) = E(e^{tX}) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx}e^{-(x-\mu)^2/2\sigma^2}\,dx$$

Haciendo $(x - \mu)/\sigma = v$ en esta integral de manera que $x = \mu + \sigma v$, $dx = \sigma\,dv$, tenemos

$$M(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ut+\sigma vt-(v^2/2)}\,dv = \frac{e^{\mu t+(\sigma^2t^2/2)}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(v-\sigma t)^2/2}\,dv$$

Ahora, haciendo $v - \sigma t = w$, encontramos

$$M(t) = e^{\mu t+(\sigma^2t^2/2)}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-w^2/2}dw\right) = e^{ut+(\sigma^2t^2/2)}$$

APROXIMACIÓN NORMAL A LA DISTRIBUCIÓN BINOMIAL

4.17. Encontrar la probabilidad de obtener entre 3 y 6 caras en 10 lanzamientos de una moneda legal, inclusive, usando *a*) la distribución binomial, *b*) la aproximación normal a la distribución binomial.

a) Sea X la variable aleatoria que da el número de caras en 10 lanzamientos de la moneda (figura 4-15). Así,

$$P(X = 3) = \binom{10}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^7 = \frac{15}{128} \qquad P(X = 4) = \binom{10}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^6 = \frac{105}{512}$$

$$P(X = 5) = \binom{10}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^5 = \frac{63}{256} \qquad P(X = 6) = \binom{10}{6}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^4 = \frac{105}{512}$$

Por tanto, la probabilidad que buscamos es

$$P(3 \le X \le 6) = \frac{15}{128} + \frac{105}{512} + \frac{63}{256} + \frac{105}{512} = \frac{99}{128} = 0.7734$$

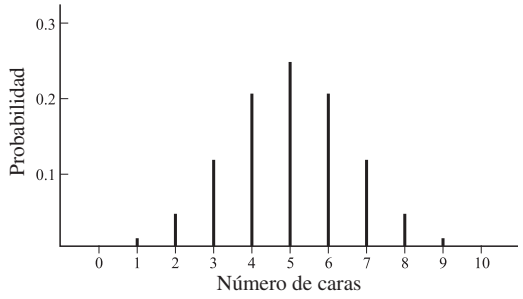


Figura 4-15

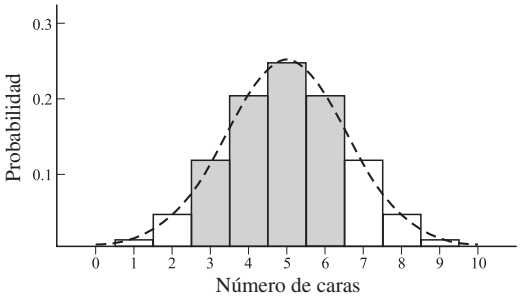


Figura 4-16

b) La distribución de probabilidad para el número de caras en 10 lanzamientos de la moneda se muestra gráficamente en las figuras 4-15 y 4-16. En esta última se tratan los datos como si fueran continuos. La probabilidad que buscamos es la suma de las áreas de los rectángulos sombreados de la figura 4-16, área que puede aproximarse por el área bajo la curva normal correspondiente, que aparece con línea punteada. Si se consideran los datos como si fueran continuos, se deduce que de 3 a 6 caras pueden considerarse como 2.5 a 6.5 caras. La media y la varianza de la distribución binomial son $\mu = np = 10\left(\frac{1}{2}\right) = 5$ y $\sigma = \sqrt{npq} = \sqrt{(10)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 1.58$.

Ahora

$$2.5 \text{ en unidades estándar} = \frac{2.5 - 5}{1.58} = -1.58$$

$$6.5 \text{ en unidades estándar} = \frac{6.5 - 5}{1.58} = 0.95$$

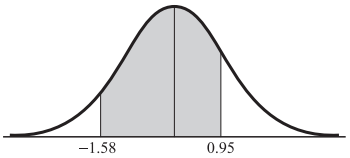


Figura 4-17

La probabilidad que buscamos (figura 4-17) = (área entre $z = -1.58$ y $z = 0.95$)
= (área entre $z = -1.58$ y $z = 0$)
+ (área entre $z = 0$ y $z = 0.95$)
= 0.4429 + 0.3289 = 0.7718

que coincide bastante bien con el verdadero valor 0.7734 que obtuvimos en el inciso *a*). La precisión es aún mayor cuando se emplean valores más grandes de *n.*

4.18. Se lanza una moneda legal 500 veces. Calcular la probabilidad de que el número de caras no difiera de 250 *a*) en más de 10, *b*) en más de 30.

$$\mu = np = (500)\left(\frac{1}{2}\right) = 250 \qquad \sigma = \sqrt{npq} = \sqrt{(500)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 11.18$$

a) La probabilidad buscada es la probabilidad de que el número de caras se encuentre entre 240 y 260 o, si se consideran estos datos como continuos, entre 239.5 y 260.5.

$$239.5 \text{ en unidades estándar} = \frac{239.5 - 250}{11.18} = -0.94 \qquad 260.5 \text{ en unidades estándar} = 0.94$$

Probabilidad que buscamos = (área bajo la curva normal entre $z = -0.94$ y $z = 0.94$)
= (el doble del área entre $z = 0$ y $z = 0.94$) = 2(0.3264) = 0.6528

b) La probabilidad que buscamos es la de que el número de caras se encuentre entre 220 y 280 o, si consideramos estos datos como continuos, entre 219.5 y 280.5.

$$219.5 \text{ en unidades estándar} = \frac{219.5 - 250}{11.18} = -2.73 \qquad 280.5 \text{ en unidades estándar} = 2.73$$

La probabilidad que buscamos = (el doble del área bajo la curva normal entre $z = 0$ y $z = 2.73$)
= 2(0.4968) = 0.9936

Concluimos que podemos confiar en que el número de caras no difiera de las esperadas (250) en más de 30. Por tanto, si se encuentra que el número *real* de caras es 280, habrá razones de peso para pensar que la moneda no es legal, es decir, está *cargada.*

4.19. Se lanza un dado 120 veces. Calcular la probabilidad de que se obtenga un 4 *a*) 18 veces o menos, *b*) 14 veces o menos, suponiendo que el dado no está cargado.

La probabilidad de obtener un 4 es $p = \frac{1}{6}$ y la probabilidad de no sacarlo es $q = \frac{5}{6}$.

a) Buscamos la probabilidad de que el número de cuatros esté entre 0 y 18. Está probabilidad es

$$\binom{120}{18}\left(\frac{1}{2}\right)^{18}\left(\frac{5}{6}\right)^{102} + \binom{120}{17}\left(\frac{1}{6}\right)^{17}\left(\frac{5}{6}\right)^{103} + \cdots + \binom{120}{0}\left(\frac{1}{6}\right)^{0}\left(\frac{5}{6}\right)^{120}$$

pero como hacer estos cálculos es muy laborioso, se usa la aproximación normal.

Considerando que los datos son continuos, concluimos que de 0 a 18 cuatros pueden tratarse como −0.5 a 18.5 cuatros. También

$$\mu = np = 120\left(\frac{1}{6}\right) = 20 \qquad \text{y} \qquad \sigma = \sqrt{npq} = \sqrt{(120)\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)} = 4.08$$

Entonces

$$-0.5 \text{ en unidades estándar} = \frac{-0.5 - 20}{4.08} = -5.02. \qquad 18.5 \text{ en unidades estándar} = -0.37$$

$$\begin{aligned}\text{Probabilidad que buscamos} &= (\text{área bajo la curva normal entre } z = -5.02 \text{ y } z = -0.37)\\ &= (\text{área entre } z = 0 \text{ y } z = -5.02)\\ &\quad - (\text{área entre } z = 0 \text{ y } z = -0.37)\\ &= 0.5 - 0.1443 = 0.3557\end{aligned}$$

b) Procedemos como en el inciso *a*), y colocamos 14 en vez de 18. Así que

$$-0.5 \text{ en unidades estándar} = -5.02 \qquad 14.5 \text{ en unidades estándar} = \frac{14.5 - 20}{4.08} = -1.35$$

$$\begin{aligned}\text{Probabilidad que buscamos} &= (\text{área bajo la curva normal entre } z = -5.02 \text{ y } z = -1.35)\\ &= (\text{área entre } z = 0 \text{ y } z = -5.02)\\ &\quad - (\text{área entre } z = 0 \text{ y } z = -1.35)\\ &= 0.5 - 0.4115 = 0.0885\end{aligned}$$

Concluimos que si se lanzara 120 veces un dado, aparecería un cuatro 14 veces, o menos, en aproximadamente un décimo de la muestra.

DISTRIBUCIÓN DE POISSON

4.20. Establecer la validez de la aproximación de Poisson a la distribución binomial.

Si X está distribuida binomialmente, entonces

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \tag{1}$$

donde $E(X) = np$. Sea $\lambda = np$ de manera que $p = \lambda/n$. Entonces, (1) se convierte en

$$\begin{aligned}P(X = x) &= \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1 - \frac{\lambda}{n}\right)^{n-x}\\ &= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!n^x}\lambda^x\left(1 - \frac{\lambda}{n}\right)^{n-x}\\ &= \frac{\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{x-1}{n}\right)}{x!}\lambda^x\left(1 - \frac{\lambda}{n}\right)^{n-x}\end{aligned}$$

Ahora, cuando $n \to \infty$,

$$\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{x-1}{n}\right) \to 1$$

mientras que

$$\left(1 - \frac{\lambda}{n}\right)^{n-x} = \left(1 - \frac{\lambda}{n}\right)^n\left(1 - \frac{\lambda}{n}\right)^{-x} \to (e^{-\lambda})(1) = e^{-\lambda}$$

usando el resultado conocido del cálculo

$$\lim_{n\to\infty}\left(1 + \frac{u}{n}\right)^n = e^u$$

Se concluye que cuando $n \to \infty$ pero λ permanece fija (es decir, $p \to 0$),

$$P(X = x) \to \frac{\lambda^x e^{-\lambda}}{x!} \tag{2}$$

que es la distribución de Poisson.

Otro método

La función generadora de momentos de la distribución binomial es

$$(q + pe^t)^n = (1 - p + pe^t)^n = [1 + p(e^t - 1)]^n \tag{3}$$

Si $\lambda = np$ de manera que $p = \lambda/n$, esta expresión se convierte en

$$\left[1 + \frac{\lambda(e^t - 1)}{n}\right]^n \tag{4}$$

Cuando $n \to \infty$ esto se aproxima a

$$e^{\lambda(e^t - 1)} \tag{5}$$

que es la función generadora de momentos de la distribución de Poisson. El resultado se basa en el teorema 3-10, página 77.

4.21. Verificar que la función límite (2) del problema 4.20 es, en realidad, una función de probabilidad.

Primero, vemos que $P(X = x) > 0$ para $x = 0, 1, \ldots$, dado que $\lambda > 0$. Segundo, tenemos

$$\sum_{x=0}^{\infty} P(X = x) = \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

con lo que termina la verificación.

4.22. Diez por ciento de las herramientas producidas mediante cierto procedimiento tienen algún defecto. Calcular la probabilidad de que en una muestra de 10 unidades elegidas al azar, exactamente 2 tengan algún defecto, empleando *a*) la distribución binomial, *b*) la aproximación de Poisson a la distribución binomial.

a) La probabilidad de que una pieza tenga algún defecto es $p = 0.1$. Sea X el número de herramientas con algún defecto en 10 herramientas. Entonces, de acuerdo con la distribución binomial,

$$P(X = 2) = \binom{10}{2}(0.1)^2(0.9)^8 = 0.1937 \text{ o bien } 0.19$$

b) Se tiene $\lambda = np = (10)(0.1) = 1$. Entonces, de acuerdo con la distribución de Poisson,

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad \text{o bien} \qquad P(X = 2) = \frac{(1)^2 e^{-1}}{2!} = 0.1839 \text{ o bien } 0.18$$

En general, la aproximación es buena si $p \leq 0.1$ y $\lambda = np \leq 5$.

4.23. Si la probabilidad de que un individuo tenga una reacción adversa a la inyección de un determinado suero es de 0.001, determinar la probabilidad de que en 2 000 individuos, *a*) exactamente 3, *b*) más de 2, tengan una reacción adversa.

Sea X el número de individuos que tienen una reacción adversa. X está distribuida de acuerdo con la distribución de Bernoulli, pero como las reacciones adversas se consideran eventos raros, puede suponerse que X está distribuida de acuerdo con la distribución de Poisson, es decir,

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad \text{donde } \lambda = np = (2\,000)(0.001) = 2$$

a)
$$P(X = 3) = \frac{2^3 e^{-2}}{3!} = 0.180$$

b)
$$\begin{aligned} P(X > 2) &= 1 - [P(X = 0) + P(X = 1) + P(X = 2)] \\ &= 1 - \left[\frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} + \frac{2^2 e^{-2}}{2!}\right] \\ &= 1 - 5e^{-2} = 0.323 \end{aligned}$$

Una evaluación exacta de estas probabilidades empleando la distribución binomial sería mucho más laboriosa.

TEOREMA DEL LÍMITE CENTRAL

4.24. Verificar el teorema del límite central para una variable aleatoria X que tiene una distribución binomial, y con ello establecer la validez de la aproximación normal a la distribución binomial.

La variable estandarizada de X es $X^* = (X - np)/\sqrt{npq}$, y la función generadora de momentos de X^* es

$$\begin{aligned}
E(e^{tX^*}) &= E(e^{t(X-np)/\sqrt{npq}}) \\
&= e^{-tnp/\sqrt{npq}}\, E(e^{tX/\sqrt{npq}}) \\
&= e^{-tnp/\sqrt{npq}} \sum_{x=0}^{n} e^{tx/\sqrt{npq}} \binom{n}{x} p^x\, q^{n-x} \\
&= e^{-tnp/\sqrt{npq}} \sum_{x=0}^{n} \binom{n}{x} (pe^{t/\sqrt{npq}})^x\, q^{n-x} \\
&= e^{-tnp/\sqrt{npq}}(q + pe^{t/\sqrt{npq}})^n \\
&= [e^{-tp/\sqrt{npq}}(q + pe^{t/\sqrt{npq}})]^n \\
&= (qe^{-tp/\sqrt{npq}} + pe^{tq/\sqrt{npq}})^n
\end{aligned}$$

Usando la expansión

$$e^u = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \cdots$$

encontramos

$$\begin{aligned}
qe^{-tp/\sqrt{npq}} + pe^{tq/\sqrt{npq}} &= q\left(1 - \frac{tp}{\sqrt{npq}} + \frac{t^2p^2}{2npq} + \cdots\right) \\
&\quad + p\left(1 + \frac{tq}{\sqrt{npq}} + \frac{t^2q^2}{2npq} + \cdots\right) \\
&= q + p + \frac{pq(p + q)t^2}{2npq} + \cdots \\
&= 1 + \frac{t^2}{2n} + \cdots
\end{aligned}$$

Por tanto,

$$E(e^{tX^*}) = \left(1 + \frac{t^2}{2n} + \cdots\right)^n$$

Sin embargo, cuando $n \to \infty$, el lado derecho tiende a $e^{t^2/2}$, que es la función generadora de momentos de la distribución normal estándar. Por tanto, el resultado que buscamos se obtiene del teorema 3-10, página 77.

4.25. Demostrar el teorema del límite central (teorema 4-2, página 112).

Para $n = 1, 2, \ldots$, tenemos $S_n = X_1 + X_2 + \cdots + X_n$. Ahora, cada una de las $X_1, X_2, \ldots, X_n$, tiene media μ y desviación estándar σ^2. Por tanto,

$$E(S_n) = E(X_1) + E(X_2) + \cdots + E(X_n) = n\mu$$

y como las X_k son independientes,

$$\mathrm{Var}(S_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n) = n\sigma^2$$

Se concluye que la variable aleatoria estandarizada correspondiente a S_n es

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

La función generadora de momentos de S_n^* es

$$\begin{aligned}
E(e^{tS_n^*}) &= E[e^{t(S_n - n\mu)/\sigma\sqrt{n}}] \\
&= E[e^{t(X_1-\mu)/\sigma\sqrt{n}} e^{t(X_2-\mu)/\sigma\sqrt{n}} \cdots e^{t(X_n-\mu)/\sigma\sqrt{n}}] \\
&= E[e^{t(X_1-\mu)/\sigma\sqrt{n}}] \cdot E[e^{t(X_2-\mu)/\sigma\sqrt{n}}] \cdots E[e^{t(X_n-\mu)/\sigma\sqrt{n}}] \\
&= \{E[e^{t(X_1-\mu)/\sigma\sqrt{n}}]\}^n
\end{aligned}$$

donde, en los últimos dos pasos, se han usado, respectivamente, los datos de que las X_k son independientes y están distribuidas de manera idéntica. Ahora, de acuerdo con la expansión de la serie de Taylor,

$$E[e^{t(X_1-\mu)/\sigma\sqrt{n}}] = E\left[1 + \frac{t(X_1-\mu)}{\sigma\sqrt{n}} + \frac{t^2(X_1-\mu)^2}{2\sigma^2 n} + \cdots\right]$$

$$= E(1) + \frac{t}{\sigma\sqrt{n}}E(X_1-\mu) + \frac{t^2}{2\sigma^2 n}E[(X_1-\mu)^2] + \cdots$$

$$= 1 + \frac{t}{\sigma\sqrt{n}}(0) + \frac{t^2}{2\sigma^2 n}(\sigma^2) + \cdots = 1 + \frac{t^2}{2n} + \cdots$$

de manera que $$E(e^{tS_n^*}) = \left(1 + \frac{t^2}{2n} + \cdots\right)^n$$

Sin embargo, el límite cuando $n \to \infty$ es $e^{t^2/2}$, que es la función generadora de momentos de la distribución normal estandarizada. Por tanto, de acuerdo con el teorema 3-10, página 80, se obtiene el resultado que buscamos.

DISTRIBUCIÓN MULTINOMIAL

4.26. Una caja contiene 5 pelotas rojas, 4 blancas y 3 azules. En forma aleatoria se extrae una de ellas, se anota su color y se reintegra a la caja. Calcular la probabilidad que de 6 pelotas extraídas de esta manera, 3 sean rojas, 2 sean blancas y 1 sea azul.

Método 1 (empleando la fórmula)

$$P(\text{roja en una extracción}) = \frac{5}{12} \qquad P(\text{blanca en una extracción}) = \frac{4}{12}$$

$$P(\text{azul en una extracción}) = \frac{3}{12}$$

Entonces, $$P(\text{3 rojas, 2 blancas y 1 azul}) = \frac{6!}{3!2!1!}\left(\frac{5}{12}\right)^3\left(\frac{4}{12}\right)^2\left(\frac{3}{12}\right)^1 = \frac{625}{5\,184}$$

Método 2

La probabilidad de extraer una de las pelotas rojas es $5/12$. Por consiguiente, la probabilidad de extraer 3 pelotas rojas es $(5/12)^3$. De manera similar, la probabilidad de extraer 2 pelotas blancas es $(4/12)^2$ y la de sacar 1 pelota azul es $(3/12)^1$. Por tanto, la probabilidad de extraer 3 rojas, 2 blancas y 1 azul, en este orden, es

$$\left(\frac{5}{12}\right)^3\left(\frac{4}{12}\right)^2\left(\frac{3}{12}\right)^1$$

Sin embargo, estas pelotas pueden extraerse en varios órdenes, y el número de estas diferentes maneras es

$$\frac{6!}{3!2!1!}$$

como se muestra en el capítulo 1. Entonces, la probabilidad que buscamos es

$$\frac{6!}{3!2!1!}\left(\frac{5}{12}\right)^3\left(\frac{4}{12}\right)^2\left(\frac{3}{12}\right)^1$$

Método 3

La probabilidad que buscamos es el término $p_r^3 p_w^2 p_b$ de la expansión multinomial de $(p_r + p_w + p_b)^6$ donde $p_r = 5/12$, $p_w = 4/12$, $p_b = 3/12$. Con esta expansión obtuvimos el anterior resultado.

DISTRIBUCIÓN HIPERGEOMÉTRICA

4.27. Una caja contiene 6 canicas azules y 4 rojas. Se realiza un experimento en el que se toma una canica en forma aleatoria y se observa de qué color es, pero la que se extrae no se devuelve a la caja. Calcular la probabilidad de que, después de cinco ensayos de este experimento, se hayan tomado 3 canicas azules.

Método 1

El número de maneras diferentes de tomar 3 canicas azules de 6 canicas de ese color es $\binom{6}{3}$. El número de maneras distintas de tomar las 2 canicas restantes de las 4 rojas es $\binom{4}{2}$. Por tanto, el número de muestras diferentes que contiene 3 canicas azules y 2 rojas es $\binom{6}{3}\binom{4}{2}$.

Ahora, el número total de maneras diferentes de tomar 5 de las 10 canicas (6 + 4) de la caja es $\binom{10}{5}$. Por tanto, la probabilidad que buscamos está dada por

$$\frac{\binom{6}{3}\binom{4}{2}}{\binom{10}{5}} = \frac{10}{21}$$

Método 2 (usando la fórmula)
Se tiene $b = 6$, $r = 4$, $n = 5$, $x = 3$. Entonces, de acuerdo con (19), página 113, la probabilidad que buscamos es

$$P(X = 3) = \frac{\binom{6}{3}\binom{4}{2}}{\binom{10}{2}}$$

DISTRIBUCIÓN UNIFORME

4.28. Demostrar que la media y la varianza de la distribución uniforme (página 113) están dados, respectivamente, por *a*) $\mu = \frac{1}{2}(a + b)$, *b*) $\sigma^2 = \frac{1}{12}(b - a)^2$.

a)
$$\mu = E(X) = \int_a^b \frac{x\,dx}{b - a} = \frac{x^2}{2(b - a)}\Big|_a^b = \frac{b^2 - a^2}{2(b - a)} = \frac{a + b}{2}$$

b) Se tiene

$$E(X^2) = \int_a^b \frac{x^2 dx}{b - a} = \frac{x^3}{3(b - a)}\Big|_a^b = \frac{b^3 - a^3}{3(b - a)} = \frac{b^2 + ab + a^2}{3}$$

Entonces la varianza está dada por

$$\sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2$$
$$= \frac{b^2 + ab + a^2}{3} - \left(\frac{a + b}{2}\right)^2 = \frac{1}{12}(b - a)^2$$

DISTRIBUCIÓN DE CAUCHY

4.29. Demostrar *a*) que la función generadora de momentos de una variable aleatoria X con la distribución de Cauchy no existe, *b*) pero que la función característica sí existe.

a) La función generadora de momentos de X es

$$E(e^{tX}) = \frac{a}{\pi}\int_{-\infty}^{\infty} \frac{e^{tx}}{x^2 + a^2}\,dx$$

la cual no existe si t es real. Este resultado se demuestra si observamos, por ejemplo, que si $x \geq 0$, $t > 0$,

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \cdots > \frac{t^2x^2}{2}$$

de manera que

$$\frac{a}{\pi}\int_{-\infty}^{\infty} \frac{e^{tx}}{x^2 + a^2}dx \geq \frac{at^2}{2\pi}\int_0^{\infty} \frac{x^2}{x^2 + a^2}dx$$

y la integral de la derecha diverge.

b) La función característica de X es

$$E(e^{tX}) = \frac{a}{\pi}\int_{-\infty}^{\infty} \frac{e^{tx}}{x^2 + a^2}\,dx$$
$$= \frac{a}{\pi}\int_{-\infty}^{\infty} \frac{\cos \omega x}{x^2 + a^2}\,dx + \frac{ai}{\pi}\int_{-\infty}^{\infty} \frac{\operatorname{sen} \omega x}{x^2 + a^2}\,dx$$
$$= \frac{2a}{\pi}\int_0^{\infty} \frac{\cos \omega x}{x^2 + a^2}\,dx$$

donde hemos usado el hecho de que los integrandos del penúltimo renglón son funciones par e impar, respectivamente. Puede demostrarse que la última integral existe y que es igual a $e^{-a\omega}$.

4.30. Sea Θ una variable aleatoria distribuida de manera uniforme en el intervalo $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$. Demostrar que $X = a \tan \Theta$, $a > 0$, sigue la distribución de Cauchy en $-\infty < x < \infty$.

La función de densidad de Θ es

$$f(\theta) = \frac{1}{\pi} \qquad -\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$$

Considerando la transformación $x = a \tan \theta$, se tiene

$$\theta = \tan^{-1}\frac{x}{a} \qquad \text{y} \qquad \frac{d\theta}{dx} = \frac{a}{x^2 + a^2} > 0$$

Entonces, de acuerdo con el teorema 2-3, página 42, la función de densidad de X es

$$g(x) = f(\theta)\left|\frac{d\theta}{dx}\right| = \frac{1}{\pi}\frac{a}{x^2 + a^2}$$

que es la distribución de Cauchy.

DISTRIBUCIÓN GAMMA

4.31. Demostrar que la media y la varianza de la distribución gamma están dadas por *a*) $\mu = \alpha\beta$, *b*) $\sigma^2 = \alpha\beta^2$.

a)
$$\mu = \int_0^\infty x\left[\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}\right]dx = \int_0^\infty \frac{x^\alpha e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}\,dx$$

Cuando $x/\beta = t$, tenemos

$$\mu = \frac{\beta^\alpha\beta}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty t^\alpha e^{-t}\,dt = \frac{\beta}{\Gamma(\alpha)}\,\Gamma(\alpha + 1) = \alpha\beta$$

b)
$$E(X^2) = \int_0^\infty x^2\left[\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}\right]dx = \int_0^\infty \frac{x^{\alpha+1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}\,dx$$

Cuando $x/\beta = t$, tenemos

$$E(X^2) = \frac{\beta^{\alpha+1}\beta}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty t^{\alpha+1}e^{-t}\,dt$$

$$= \frac{\beta^2}{\Gamma(\alpha)}\,\Gamma(\alpha + 2) = \beta^2(\alpha + 1)\alpha$$

donde $\Gamma(\alpha + 2) = (\alpha + 1)\Gamma(\alpha + 1) = (\alpha + 1)\alpha\Gamma(\alpha)$. Por tanto,

$$\sigma^2 = E(X^2) - \mu^2 = \beta^2(\alpha + 1)\alpha - (\alpha\beta)^2 = \alpha\beta^2$$

DISTRIBUCIÓN BETA

4.32. Encontrar la media de la distribución beta.

$$\mu = E(X) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x[x^{\alpha-1}(1 - x)^{\beta-1}]\,dx$$

$$= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^\alpha(1 - x)^{\beta-1}\,dx$$

$$= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}\,\frac{\Gamma(\alpha + 1)\Gamma(\beta)}{\Gamma(\alpha + 1 + \beta)}$$

$$= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}\,\frac{\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha + \beta)\Gamma(\alpha + \beta)} = \frac{\alpha}{\alpha + \beta}$$

4.33. Encontrar la varianza de la distribución beta.

El segundo momento respecto al origen es

$$\begin{aligned} E(X^2) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^2[x^{\alpha-1}(1-x)^{\beta-1}]\,dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha+1}(1-x)^{\beta-1}\,dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+2+\beta)} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{(\alpha+1)\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)} \\ &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \end{aligned}$$

Después, empleando el problema 4.32, la varianza es

$$\sigma^2 = E(X^2) - [E(X)]^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

DISTRIBUCIÓN JI CUADRADA

4.34. Demostrar que la función generadora de momentos de la variable aleatoria X, que tiene distribución ji cuadrada con ν grados de libertad, es $M(t) = (1-2t)^{-\nu/2}$.

$$\begin{aligned} M(t) = E(e^{tX}) &= \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\int_0^\infty e^{tx}x^{(\nu-2)/2}e^{-x/2}\,dx \\ &= \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\int_0^\infty x^{(\nu-2)/2}e^{-(1-2t)x/2}\,dx \end{aligned}$$

Haciendo $(1-2t)x/2 = u$ en la última integral encontramos

$$\begin{aligned} M(t) &= \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\int_0^\infty \left(\frac{2u}{1-2t}\right)^{(\nu-2)/2} e^{-u}\frac{2\,du}{1-2t} \\ &= \frac{(1-2t)^{-\nu/2}}{\Gamma(\nu/2)}\int_0^\infty u^{(\nu/2)-1}e^{-u}\,du = (1-2t)^{-\nu/2} \end{aligned}$$

4.35. Sean X_1 y X_2 variables aleatorias independientes que tienen la distribución ji cuadrada con ν_1 y ν_2 grados de libertad, respectivamente. *a*) Demostrar que la función generadora de momentos de $Z = X_1 + X_2$ es $(1-2t)^{-(\nu_1+\nu_2)/2}$, y *b*) demostrar que Z tiene distribución ji cuadrada con $\nu_1 + \nu_2$ grados de libertad.

a) La función generadora de momentos de $Z = X_1 + X_2$ es

$$M(t) = E[e^{t(X_1+X_2)}] = E(e^{tX_1})E(e^{tX_2}) = (1-2t)^{-\nu_1/2}(1-2t)^{-\nu_2/2} = (1-2t)^{-(\nu_1+\nu_2)/2}$$

con base en el problema 4.34.

b) De acuerdo con el problema 4.34 observamos que una distribución cuya función generadora de momentos es $(1-2t)^{-(\nu_1+\nu_2)/2}$, es la distribución ji cuadrada con $\nu_1 + \nu_2$ grados de libertad. Ésta debe ser la distribución de Z, de acuerdo con el teorema 3-10, página 77.

Por generalización de los resultados anteriores obtenemos la demostración del teorema 4-4, página 115.

4.36. Sea X una variable aleatoria distribuida normalmente con media 0 y varianza 1. Demostrar que X^2 tiene la distribución ji cuadrada con 1 grado de libertad.

Se quiere hallar la distribución de $Y = X^2$ dada la distribución normal estándar de X. Como la correspondencia entre X y Y no es uno a uno, no podemos emplear el teorema 2-3 tal como está, pero lo haremos como sigue.

Para $y < 0$ es claro que $P(Y \le y) = 0$. Para $y \ge 0$, tenemos

$$P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le +\sqrt{y})$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\sqrt{y}}^{+\sqrt{y}} e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}}\int_{0}^{+\sqrt{y}} e^{-x^2/2}\,dx$$

donde en el último paso se usa el hecho de que la función de densidad normal estándar es par. Si en la última integral hacemos el cambio de variable $x = +\sqrt{t}$ se obtiene

$$P(Y \le y) = \frac{1}{\sqrt{2\pi}}\int_{0}^{y} t^{-1/2}e^{-t/2}\,dt$$

Pero ésta es una distribución ji cuadrada con 1 grado de libertad, como puede verse para $\nu = 1$ en (39), página 115, y usamos el hecho de que $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

4.37. Demostrar el teorema 4-3, página 115, para $\nu = 2$.

De acuerdo con el problema 4.36 podemos ver que si X_1 y X_2 están distribuidas normalmente con media 0 y varianza 1, entonces X_1^2 y X_2^2 tienen, cada una, distribución ji cuadrada con 1 grado de libertad. Entonces, de acuerdo con el problema 4.35*b*), se ve que $Z = X_1^2 + X_2^2$ tiene distribución ji cuadrada con $1 + 1 = 2$ grados de libertad si X_1 y X_2 son independientes. El resultado general para todo entero positivo ν se obtiene de la misma manera.

4.38. En la figura 4-18 se muestra la gráfica de la distribución ji cuadrada con 5 grados de libertad. (Vea las notas sobre la notación en la página 115.) Encontrar los valores χ_1^2, χ_2^2 para los que

a) área sombreada de la derecha = 0.05,

b) área sombreada total = 0.05

c) área sombreada de la izquierda = 0.10,

d) área sombreada de la derecha = 0.01.

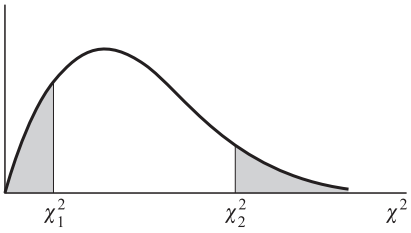


Figura 4-18

a) Si el área sombreada de la derecha es 0.05, entonces el área de la izquierda de χ_2^2 es $(1 - 0.05) = 0.95$, y χ_2^2 representa el percentil 95, $\chi^2_{0.95}$.

En el apéndice E se desciende por la columna cuyo encabezado es ν hasta llegar a la entrada 5. Después se avanza hacia la derecha hasta la columna cuyo encabezado es $\chi^2_{0.95}$. El resultado, 11.1, es el valor que se busca de χ^2.

b) Como la distribución no es simétrica, hay muchos valores para los que el área sombreada total = 0.05. Por ejemplo, el área sombreada de la derecha puede ser 0.04 y el área sombreada de la izquierda 0.01. Sin embargo, se acostumbra, a menos que se especifique otra cosa, elegir las dos áreas iguales. En este caso, cada área = 0.025.

Si el área sombreada de la derecha es 0.025, entonces el área de la izquierda de χ_2^2 es $1 - 0.025 = 0.975$ y χ_2^2 representa el percentil 97.5, $\chi^2_{0.975}$, que de acuerdo con el apéndice E es 12.8.

De manera similar, si el área sombreada de la izquierda es 0.025, el área de la izquierda de χ_1^2 es 0.025 y χ_1^2 representa el percentil 2.5, $\chi^2_{0.025}$, que es igual a 0.831.

Por tanto, los valores son 0.831 y 12.8.

c) Si el área sombreada de la izquierda es 0.10, χ_1^2 representa el percentil 10, $\chi^2_{0.10}$, que es igual a 1.61.

d) Si el área sombreada de la derecha es 0.01, el área de la izquierda de χ_2^2 es 0.99 y χ_2^2 representa el percentil 99, $\chi^2_{0.99}$, que es a igual a 15.1.

4.39. Determinar los valores de χ^2 para los que el área de la cola derecha de la distribución χ^2 es 0.05, si el número de grados de libertad ν es igual a *a*) 15, *b*) 21, *c*) 50.

Usando la tabla del apéndice E, en la columna cuyo encabezado es $\chi^2_{0.95}$ leemos los valores: *a*) 25.0 que corresponde a $\nu = 15$; *b*) 32.7 que corresponde a $\nu = 21$; *c*) 67.5 que corresponde a $\nu = 50$.

4.40. Encontrar la mediana de χ^2 que corresponde a *a*) 9, *b*) 28 y *c*) 40 grados de libertad.

Usando el apéndice E, en la columna cuyo encabezado es $\chi^2_{0.50}$ (ya que la mediana es el percentil 50) vemos los valores: *a*) 8.34 que corresponde a $\nu = 9$; *b*) 27.3 que corresponde a $\nu = 28$; *c*) 39.3 que corresponde a $\nu = 40$.

Observemos que los valores de la mediana son casi iguales al número de grados de libertad. En efecto, para $\nu > 10$, los valores de la mediana son iguales a $\nu - 0.7$, como puede verse en la tabla.

4.41. Calcular $\chi^2_{0.95}$, para *a*) $\nu = 50$, *b*) $\nu = 100$ grados de libertad.

Para ν mayor a 30, puede usarse el hecho de que $(\sqrt{2\chi^2} - \sqrt{2v - 1})$ está distribuida casi normalmente con media cero y varianza uno. Entonces, si z_p es el percentil $(100p)$ de la distribución normal estándar, podemos escribir, con un alto grado de aproximación,

$$\sqrt{2\chi_p^2} - \sqrt{2\nu - 1} = z_p \qquad \text{o} \qquad \sqrt{2\chi_p^2} = z_p + \sqrt{2\nu - 1}$$

de donde

$$\chi_p^2 = \tfrac{1}{2}(z_p + \sqrt{2\nu - 1})^2$$

a) Si $\nu = 50$, $\chi^2_{0.95} = \frac{1}{2}(z_{0.95} + \sqrt{2(50) - 1})^2 = \frac{1}{2}(1.64 + \sqrt{99})^2 = 69.2$, que coincide muy bien con el valor 67.5 dado en el apéndice E.

b) Si $\nu = 100$, $\chi^2_{0.95} = \frac{1}{2}(z_{0.95} + \sqrt{2(100) - 1})^2 = \frac{1}{2}(1.64 + \sqrt{199})^2 = 124.0$ (valor verdadero = 124.3).

DISTRIBUCIÓN *t* DE STUDENT

4.42. Demostrar el teorema 4-6, página 116.

Como Y está distribuida normalmente con media 0 y varianza 1, su función de densidad es

$$\frac{1}{\sqrt{2\pi}}\, e^{-y^2/2} \tag{1}$$

Como Z sigue la distribución ji cuadrada con ν grados de libertad, su función de densidad es

$$\frac{1}{2^{\nu/2}\Gamma(\nu/2)}\, z^{(\nu/2)-1} e^{-z/2} \qquad z > 0 \tag{2}$$

Dado que Y y Z son independientes, su función de densidad conjunta es el producto de (1) y (2), es decir,

$$\frac{1}{\sqrt{2\pi}\,2^{\nu/2}\Gamma(\nu/2)}\, z^{(\nu/2)-1}\, e^{-(y^2+z)/2}$$

para $-\infty < y < +\infty$, $z > 0$.

La función de distribución de $T = Y/\sqrt{Z/\nu}$ es

$$F(x) = P(T \le x) = P(Y \le x\sqrt{Z/\nu})$$

$$= \frac{1}{\sqrt{2\pi}2^{\nu/2}\Gamma(\nu/2)} \iint_{\mathcal{R}} z^{(\nu/2)-1} e^{-(y^2+z)/2}\, dy\, dz$$

donde integramos sobre la región $\mathcal{R}$ del plano yz para el que $y \le x\sqrt{z/\nu}$. Primero se deja fija z e integramos respecto a y desde $-\infty$ hasta $x\sqrt{z/\nu}$. Después se integra respecto a z desde 0 hasta ∞. Entonces tenemos

$$F(x) = \frac{1}{\sqrt{2\pi}2^{\nu/2}\Gamma(\nu/2)} \int_{z=0}^{\infty} z^{(\nu/2)-1} e^{-z/2} \left[\int_{y=-\infty}^{x\sqrt{z/\nu}} e^{-y^2/2}\, dy \right] dz$$

Hacemos $y = u\sqrt{z/\nu}$ en la integral entre corchetes, encontramos

$$F(x) = \frac{1}{\sqrt{2\pi}\,2^{\nu/2}\Gamma(\nu/2)}\int_{z=0}^{\infty}\int_{u=-\infty}^{\infty} z^{(\nu/2)-1}e^{-z/2}\sqrt{z/\nu}\,e^{-u^2z/2\nu}\,du\,dz$$

$$= \frac{1}{\sqrt{2\pi\nu}\,2^{\nu/2}\Gamma(\nu/2)}\int_{u=-\infty}^{x}\left[\int_{z=0}^{\infty} z^{(\nu-1)/2}\,e^{-(z/2)[1+(u^2/\nu)]}\,dz\right]du$$

Hacemos $w = \frac{z}{2}\left(1 + \frac{u^2}{\nu}\right)$ podemos escribir

$$F(x) = \frac{1}{\sqrt{2\pi\nu}\,2^{\nu/2}\Gamma(\nu/2)} \cdot 2^{(\nu+1)/2}\int_{u=-\infty}^{x}\left[\int_{w=0}^{\infty}\frac{w^{(\nu-1)/2}e^{-w}}{(1 + u^2/\nu)^{(\nu+1)/2}}\,dw\right]du$$

$$= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\int_{u=-\infty}^{x}\frac{du}{(1 + u^2/\nu)^{(\nu+1)/2}}$$

como deseábamos.

4.43. En la figura 4-19 se muestra la gráfica de la distribución t de Student con 9 grados de libertad. Encontrar el valor de t_1 para el que

a) el área sombreada de la derecha = 0.05,

b) toda el área sombreada = 0.05,

c) toda el área no sombreada = 0.99,

d) el área sombreada de la izquierda = 0.01,

e) el área a la izquierda de t_1 es 0.90.

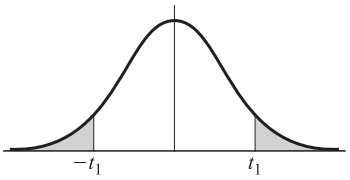


Figura 4-19

a) Si el área sombreada de la derecha es 0.05, entonces el área a la izquierda de t_1 es $(1 - 0.05) = 0.95$ y t_1 representa el percentil 95, $t_{0.95}$.

En la tabla del apéndice D bajamos por la columna cuyo encabezado es ν hasta llegar a la entrada 9. Después avanzamos a la derecha hasta la columna cuyo encabezado es $t_{0.95}$. El resultado 1.83 es el valor que buscamos de t.

b) Si toda el área sombreada es 0.05, entonces, por simetría, el área sombreada de la derecha es 0.025. Por tanto, el área de la izquierda de t_1 es $(1 - 0.025) = 0.975$ y t_1 representa el percentil 97.5, $t_{0.975}$. En el apéndice D encontramos 2.26 como valor que buscamos de t.

c) Si toda el área no sombreada es 0.99, entonces toda el área sombreada es $(1 - 0.99) = 0.01$ y el área sombreada de la derecha es $0.01/2 = 0.005$. En la tabla encontramos $t_{0.995} = 3.25$.

d) Si el área sombreada de la izquierda es 0.01, entonces, por simetría, el área sombreada de la derecha es 0.01. En la tabla $t_{0.99} = 2.82$. Por tanto, el valor de t para que el área sombreada de la izquierda sea 0.01 es -2.82.

e) Si el área sombreada de la izquierda de t_1 es 0.90, entonces t_1 corresponde al percentil 90, $t_{0.90}$, que de acuerdo con la tabla es igual a 1.38.

4.44. Encontrar los valores de t para los que el área en la cola derecha de la distribución t es 0.05 si el número de grados de libertad ν es igual a *a*) 16, *b*) 27, *c*) 200.

En el apéndice D, en la columna cuyo encabezado es $t_{0.95}$, encontramos los valores: *a*) 1.75 que corresponde a $\nu = 16$; *b*) 1.70 que corresponde a $\nu = 27$; *c*) 1.645 que corresponde a $\nu = 200$. (Este último valor es el que se obtendría si se emplea la curva normal. En el apéndice D este valor corresponde a la entrada en la última columna marcada con ∞.)

Distribución *F*

4.45. Demostrar el teorema 4-7.

La función de densidad conjunta de V_1 y V_2 está dada por

$$f(v_1, v_2) = \left(\frac{1}{2^{\nu_1/2}\Gamma(\nu_1/2)}\, v_1^{(\nu_1/2)-1} e^{-v_1/2}\right)\left(\frac{1}{2^{\nu_2/2}\Gamma(\nu_2/2)}\, v_2^{(\nu_2/2)-1} e^{-v_2/2}\right)$$

$$= \frac{1}{2^{(\nu_1+\nu_2)/2}\Gamma(\nu_1/2)\Gamma(\nu_2/2)}\, v_1^{(\nu_1/2)-1} v_2^{(\nu_2/2)-1} e^{-(v_1+v_2)/2}$$

si $v_1 > 0$, $v_2 > 0$ y 0 si no es así. Se hace la transformación

$$u = \frac{v_1/\nu_1}{v_2/\nu_2} = \frac{\nu_2 v_1}{\nu_1 v_2}, \qquad w = v_2 \quad \text{o bien} \quad v_1 = \frac{\nu_1 u w}{\nu_2} \qquad v_2 = w$$

En consecuencia, el jacobiano es

$$\frac{\partial(v_1, v_2)}{\partial(u, w)} = \begin{vmatrix} \partial v_1/\partial u & \partial v_1/\partial w \\ \partial v_2/\partial u & \partial v_2/\partial w \end{vmatrix} = \begin{vmatrix} \nu_1 w/\nu_2 & \nu_1 u/\nu_2 \\ 0 & 1 \end{vmatrix} = \frac{\nu_1 w}{\nu_2}$$

Denotando la densidad como función de u y w, $g(u, w)$, tenemos

$$g(u, w) = \frac{1}{2^{(\nu_1+\nu_2)/2}\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_1 u w}{\nu_2}\right)^{(\nu_1/2)-1} w^{(\nu_2/2)-1} e^{-[1+(\nu_1 u/\nu_2)](w/2)}\, \frac{\nu_1 w}{\nu_2}$$

si $u > 0$, $w > 0$ y 0 si no es así.

Ahora, la función de densidad (marginal) de U puede calcularse si se integra respecto a w desde 0 hasta ∞, es decir,

$$h(u) = \frac{(\nu_1/\nu_2)^{\nu_1/2} u^{(\nu_1/2)-1}}{2^{(\nu_1+\nu_2)/2}\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \int_0^\infty w^{[(\nu_1+\nu_2)/2]-1} e^{-[1+(\nu_1 u/\nu_2)](w/2)}\, dw$$

si $u > 0$ y 0 si $u \leq 0$. Sin embargo, de acuerdo con (15), apéndice A,

$$\int_0^\infty w^{p-1} e^{-aw}\, dw = \frac{\Gamma(p)}{a^p}$$

Por tanto, tenemos

$$h(u) = \frac{(\nu_1/\nu_2)^{\nu_1/2} u^{(\nu_1/2)-1} \Gamma\left(\dfrac{\nu_1 + \nu_2}{2}\right)}{2^{(\nu_1+\nu_2)/2}\Gamma(\nu_1/2)\Gamma(\nu_2/2)\left[\dfrac{1}{2}\left(1 + \dfrac{\nu_1 u}{\nu_2}\right)\right]^{(\nu_1+\nu_2)/2}}$$

$$= \frac{\Gamma\left(\dfrac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)} \nu_1^{\nu_1/2} \nu_2^{\nu_2/2} u^{(\nu_1/2)-1} (\nu_2 + \nu_1 u)^{-(\nu_1+\nu_2)/2}$$

si $u > 0$ y 0 si $u \leq 0$, que es el resultado que buscamos.

4.46. Demuestre que la distribución F es unimodal en el valor $\left(\frac{\nu_1 - 2}{\nu_1}\right)\left(\frac{\nu_2}{\nu_2 + 2}\right)$ si $\nu_1 > 2$.

La moda se localiza en el valor máximo de la función de densidad. Aparte de una constante, la función de densidad de la distribución F es

$$u^{(\nu_1/2)-1}(\nu_2 + \nu_1 u)^{-(\nu_1+\nu_2)/2}$$

Si esta función tiene un máximo relativo, éste se presentará donde la derivada es cero, es decir,

$$\left(\frac{\nu_1}{2} - 1\right)u^{(\nu_1/2)-2}(\nu_2 + \nu_1 u)^{-(\nu_1+\nu_2)/2} - u^{(\nu_1/2)-1}\nu_1\left(\frac{\nu_1 + \nu_2}{2}\right)(\nu_2 + \nu_1 u)^{-[(\nu_1+\nu_2)/2]-1} = 0$$

Dividiendo entre $u^{(\nu_1/2)-2}(\nu_2 + \nu_1 u)^{-[(\nu_1+\nu_2)/2]-1}$, encontramos

$$\left(\frac{\nu_1}{2} - 1\right)(\nu_2 + \nu_1 u) - u\nu_1\left(\frac{\nu_1 + \nu_2}{2}\right) = 0 \quad \text{o bien} \quad u = \left(\frac{\nu_1 - 2}{\nu_1}\right)\left(\frac{\nu_2}{\nu_2 + 2}\right)$$

Con el criterio de la segunda derivada, se puede demostrar que, en efecto, esto da el máximo.

4.47. Con base en la tabla de la distribución F que se presenta en el apéndice F, calcular *a)* $F_{0.95,10,15}$, *b)* $F_{0.99,15,9}$, *c)* $F_{0.05,8,30}$, *d)* $F_{0.01,15,9}$.

a) En el apéndice F, para $\nu_1 = 10$, $\nu_1 = 15$, encontramos $F_{0.95,10,15} = 2.54$.

b) En el apéndice F, para $\nu_1 = 15$, $\nu_2 = 9$, encontramos $F_{0.99,15,9} = 4.96$.

c) De acuerdo con el teorema 4-8, página 117, $F_{0.05,8,30} = \dfrac{1}{F_{0.95,30,8}} = \dfrac{1}{3.08} = 0.325$.

d) De acuerdo con el teorema 4-8, página 117, $F_{0.01,15,9} = \dfrac{1}{F_{0.99,9,15}} = \dfrac{1}{3.89} = 0.257$.

RELACIÓN ENTRE LAS DISTRIBUCIONES F, χ^2 Y t

4.48. Verificar que *a)* $F_{0.95} = t^2_{0.975}$, *b)* $F_{0.99} = t^2_{0.995}$.

a) Comparar las entradas de la primera columna de $F_{0.95}$, de la tabla del apéndice F, con las de la distribución t bajo $t_{0.975}$. Observamos que

$$161 = (12.71)^2, \qquad 18.5 = (4.30)^2, \qquad 10.1 = (3.18)^2, \qquad 7.71 = (2.78)^2, \qquad \text{etcétera.}$$

b) Comparar las entradas de la primera columna de $F_{0.99}$, de la tabla del apéndice F, con las de la distribución t bajo $t_{0.995}$. Observamos que

$$4\,050 = (63.66)^2, \qquad 98.5 = (9.92)^2, \qquad 34.1 = (5.84)^2, \qquad 21.2 = (4.60)^2, \qquad \text{etcétera.}$$

4.49. Demostrar el teorema 4-9 de la página 117, el cual puede enunciarse brevemente como

$$F_{1-p} = t^2_{1-(p/2)}$$

y, por tanto, generalizar los resultados del problema 4.48.

En la función de densidad de la distribución F [(45), página 116] sean $\nu_1 = 1$, $\nu_2 = \nu$. Entonces,

$$\begin{aligned} f(u) &= \frac{\Gamma\left(\frac{\nu + 1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}\nu^{\nu/2}u^{-1/2}(\nu + u)^{-(\nu+1)/2} \\ &= \frac{\Gamma\left(\frac{\nu + 1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\nu^{\nu/2}u^{-1/2}\nu^{-(\nu+1)/2}\left(1 + \frac{u}{\nu}\right)^{-(\nu+1)/2} \\ &= \frac{\Gamma\left(\frac{\nu + 1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}u^{-1/2}\left(1 + \frac{u}{\nu}\right)^{-(\nu+1)/2} \end{aligned}$$

para $u > 0$, y $f(u) = 0$ para $u \le 0$. Ahora, de acuerdo con la definición de un valor percentil, F_{1-p} es el número tal que $P(U \le F_{1-p}) = 1 - p$. Por tanto,

$$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \int_0^{F_{1-p}} u^{-1/2}\left(1 + \frac{u}{\nu}\right)^{-(\nu+1)/2} du = 1 - p$$

En la integral se hace el cambio de variable $t = +\sqrt{u}$:

$$2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\sqrt{F_{1-p}}} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2} dt = 1 - p$$

Cuando se compara con (42), página 115, se ve que el lado izquierdo de la última ecuación es igual a

$$2 \cdot P(0 < T \le +\sqrt{F_{1-p}})$$

donde T es una variable aleatoria con distribución t de Student con ν grados de libertad. Por tanto,

$$\begin{aligned} \frac{1-p}{2} &= P(0 < T \le +\sqrt{F_{1-p}}) \\ &= P(T \le +\sqrt{F_{1-p}}) - P(T \le 0) \\ &= P(T \le +\sqrt{F_{1-p}}) - \tfrac{1}{2} \end{aligned}$$

donde se usa la simetría de la distribución t. Despejando, tenemos

$$P(T \le +\sqrt{F_{1-p}}) = 1 - \frac{p}{2}$$

Sin embargo, por definición, $t_{1-(p/2)}$ es el número tal que

$$P(T \le t_{1-(p/2)}) = 1 - \frac{p}{2}$$

y este número está determinado de manera única, ya que la función de densidad de la distribución t es estrictamente positiva. Por tanto,

$$+\sqrt{F_{1-p}} = t_{1-(p/2)} \qquad \text{o bien} \qquad F_{1-p} = t^2_{1-(p/2)}$$

que era lo que deseábamos demostrar.

4.50. Verificar el teorema 4-10 de la página 117, para *a*) $p = 0.95$, *b*) $p = 0.99$.

a) Al comparar las entradas del último renglón de $F_{0.95}$ en la tabla del apéndice F (que corresponden a $\nu_2 = \infty$) con las entradas bajo $\chi^2_{0.95}$ en el apéndice E. Se observa que

$$3.84 = \frac{3.84}{1}, \qquad 3.00 = \frac{5.99}{2}, \qquad 2.60 = \frac{7.81}{3}, \qquad 2.37 = \frac{9.49}{4}, \qquad 2.21 = \frac{11.1}{5}, \qquad \text{etc.}$$

con lo que obtenemos la verificación deseada.

b) Se comparan las entradas del último renglón de $F_{0.99}$ de la tabla del apéndice F (que corresponden a $\nu_2 = \infty$) con las entradas bajo $\chi^2_{0.99}$ del apéndice E. Se observa que

$$6.63 = \frac{6.63}{1}, \qquad 4.61 = \frac{9.21}{2}, \qquad 3.78 = \frac{11.3}{3}, \qquad 3.32 = \frac{13.3}{4}, \qquad 3.02 = \frac{15.1}{5}, \qquad \text{etc.}$$

con lo que se obtiene la verificación que buscamos.

La prueba general del teorema 4-10 la obtenemos al hacer que $\nu_2 \to \infty$ en la distribución F de la página 116.

DISTRIBUCIÓN NORMAL BIVARIADA

4.51. Se supone que X y Y son variables aleatorias cuya función de densidad conjunta es la distribución normal bivariada. Demostrar que X y Y son independientes si y sólo si su coeficiente de correlación es cero.

Si el coeficiente de correlación es $\rho = 0$, entonces la función de densidad normal bivariada (49) de la página 117 se transforma en

$$f(x, y) = \left[\frac{1}{\sigma_1\sqrt{2\pi}}\, e^{-(x-\mu_1)^2/2\sigma_1^2}\right]\left[\frac{1}{\sigma_2\sqrt{2\pi}}\, e^{-(y-\mu_2)^2/2\sigma_2^2}\right]$$

y como éste es un producto de una función sólo de x por una función sólo de y para todos los valores de x y y, concluimos que X y Y son independientes.

De manera inversa, si X y Y son independientes, $f(x, y)$ dada por (49) debe ser el producto de una función sólo de x por una función sólo de y para todos los valores de x y de y. Esto es posible sólo si $\rho = 0$.

DIVERSAS DISTRIBUCIONES

4.52. Calcular la probabilidad de que en lanzamientos sucesivos de un dado no cargado, se obtenga un 3 por primera vez en el quinto lanzamiento.

Método 1

La probabilidad de no obtener un 3 en el primer lanzamiento es $5/6$; de manera similar, la probabilidad de no obtener un 3 en el segundo lanzamiento es $5/6$, etc. Entonces, la probabilidad de no obtener un 3 en los cuatro primeros lanzamientos es $(5/6)\,(5/6)\,(5/6)\,(5/6) = (5/6)^4$. Por tanto, como la probabilidad de obtener un 3 en el quinto lanzamiento es $1/6$, la probabilidad que se busca es

$$\left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right) = \frac{625}{7\,776}$$

Método 2 (con base en la fórmula)

Si empleamos la distribución geométrica de la página 117, con $p = 1/6$, $q = 5/6$, $x = 5$, se observa que la probabilidad que buscamos es

$$\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^4 = \frac{625}{7\,776}$$

4.53. Verificar las expresiones dadas para *a*) la media, *b*) la varianza, de la distribución de Weibull, página 118.

a)
$$\begin{aligned}\mu = E(X) &= \int_0^\infty abx^b e^{-ax^b}\, dx \\ &= \frac{ab}{a^{1/b}}\int_0^\infty \left(\frac{u}{a}\right) e^{-u}\,\frac{1}{b}\,u^{(1/b)-1}\, du \\ &= a^{-1/b}\int_0^\infty u^{1/b} e^{-u}\, du \\ &= a^{-1/b}\Gamma\left(1 + \frac{1}{b}\right)\end{aligned}$$

donde sustituimos $u = ax^b$ para evaluar la integral

b)
$$\begin{aligned}E(X^2) &= \int_0^\infty abx^{b+1} e^{-ax^b} dx \\ &= \frac{ab}{a^{1/b}}\int_0^\infty \left(\frac{u}{a}\right)^{1+(1/b)} e^{-u}\,\frac{1}{b}\,u^{(1/b)-1}\, du \\ &= a^{-2/b}\int_0^\infty u^{2/b}\, e^{-u}\, du \\ &= a^{-2/b}\Gamma\left(1 + \frac{2}{b}\right)\end{aligned}$$

En consecuencia,

$$\begin{aligned}\sigma^2 &= E[(X - \mu)^2] = E(X^2) - \mu^2 \\ &= a^{-2/b}\left[\Gamma\left(1 + \frac{2}{b}\right) - \Gamma^2\left(1 + \frac{1}{b}\right)\right]\end{aligned}$$

PROBLEMAS DIVERSOS

4.54. La probabilidad de que un estudiante que ingresa a la universidad se gradúe es 0.4. Determinar la probabilidad que de 5 estudiantes *a*) ninguno, *b*) 1, *c*) por lo menos 1, se gradúe.

a) $P(\text{ninguno se gradúe}) = {}_5C_0(0.4)^0(0.6)^5 = 0.07776$, o sea, aproximadamente, 0.08.

b) $P(\text{1 se gradúe}) = {}_5C_1(0.4)^1(0.6)^4 = 0.2592$, o sea, aproximadamente, 0.26.

c) $P(\text{por lo menos 1 se gradúe}) = 1 - P(\text{ninguno se gradúe}) = 0.92224$, o sea, aproximadamente, 0.92.

4.55. ¿Cuál es la probabilidad de obtener una suma total de 9 *a*) dos veces, *b*) por lo menos 2 veces, en seis lanzamientos de un par de dados no cargados?

Cada una de las 6 maneras en que puede caer el primer dado se puede relacionar con cada una de las 6 maneras en que puede caer el segundo dado, de modo que hay $6 \cdot 6 = 36$ maneras en las que pueden caer los dos dados. Éstas son: 1 en el primer dado y 1 en el segundo dado, 1 en el primer dado y 2 en el segundo dado, etc., que se denotan (1, 1), (1, 2), etcétera.

De estas 36 maneras, todas igualmente posibles si el dado no está cargado, un total de 9 se obtiene en 4 casos: (3, 6), (4, 5), (5, 4), (6, 3). Entonces, la probabilidad de obtener en total 9 en un solo lanzamiento de un par de dados es $p = 4/36 = 1/9$ y la probabilidad de no obtener en total 9 en un solo lanzamiento de un par de dados es $q = 1 - p = 8/9$.

a) $$P(\text{2 nueves en 6 lanzamientos}) = {}_6C_2\left(\frac{1}{9}\right)^2\left(\frac{8}{9}\right)^{6-2} = \frac{61\,440}{531\,441}$$

b) $$P(\text{por lo menos 2 nueves}) = P(\text{2 nueves}) + P(\text{3 nueves}) + P(\text{4 nueves}) + P(\text{5 nueves}) + P(\text{6 nueves})$$

$$= {}_6C_2\left(\frac{1}{9}\right)^2\left(\frac{8}{9}\right)^4 + {}_6C_3\left(\frac{1}{9}\right)^3\left(\frac{8}{9}\right)^3 + {}_6C_4\left(\frac{1}{9}\right)^4\left(\frac{8}{9}\right)^2 + {}_6C_5\left(\frac{1}{9}\right)^5\frac{8}{9} + {}_6C_6\left(\frac{1}{9}\right)^2$$

$$= \frac{61\,440}{531\,441} + \frac{10\,240}{531\,441} + \frac{960}{531\,441} + \frac{48}{531\,441} + \frac{1}{531\,441} = \frac{72\,689}{531\,441}$$

Otro método

$$P(\text{por lo menos 2 nueves}) = 1 - P(\text{0 nueves}) - P(\text{1 nueve})$$

$$= 1 - {}_6C_0\left(\frac{1}{9}\right)^0\left(\frac{8}{9}\right)^6 - {}_6C_1\left(\frac{1}{9}\right)^1\left(\frac{8}{9}\right)^5 = \frac{72\,689}{531\,441}$$

4.56. Si la probabilidad de que un tornillo tenga algún defecto es 0.1, calcular *a*) la media, *b*) la desviación estándar de la distribución de los tornillos con algún defecto de un total de 400.

a) Media $= np = 400(0.1) = 40$, es decir, puede *esperarse* que haya 40 tornillos defectuosos.

b) Varianza $= npq = 400(0.1)(0.9) = 36$. Por tanto, la desviación estándar es $\sqrt{36} = 6$.

4.57. Encontrar los coeficientes de *a*) sesgo, *b*) curtosis de la distribución del problema 4.56.

a) $$\text{Coeficiente de sesgo} = \frac{q - p}{\sqrt{npq}} = \frac{0.9 - 0.1}{6} = 0.133$$

Dado que este coeficiente es positivo, la distribución es sesgada a la derecha.

b) $$\text{Coeficiente de curtosis} = 3 + \frac{1 - 6pq}{npq} = 3 + \frac{1 - 6(0.1)(0.9)}{36} = 3.01$$

Esta distribución es ligeramente más afilada que la distribución normal.

4.58. Las calificaciones en un examen de biología fueron 0, 1, 2, . . . , 10 puntos, de acuerdo con la cantidad de las 10 preguntas contestadas correctamente. La calificación promedio fue 6.7 y la desviación estándar fue 1.2. Si se supone que estas calificaciones están distribuidas normalmente, determinar *a*) el porcentaje de estudiantes que obtuvieron 6, *b*) la calificación máxima de 10% inferior de la clase, *c*) la calificación mínima de 10% superior de la clase.

a) Para aplicar la distribución normal a datos discretos, es necesario tratar los datos como si fueran continuos. Así, una calificación de 6 se considera como de 5.5 a 6.5. Vea la figura 4-20.

$$5.5 \text{ en unidades estándar} = (5.5 - 6.7)/1.2 = -1.0$$

$$6.5 \text{ en unidades estándar} = (6.5 - 6.7)/1.2 = -0.17$$

$$\begin{aligned}\text{Proporción buscada} &= \text{área entre } z = -1 \text{ y } z = -0.17\\ &= (\text{área entre } z = -1 \text{ y } z = 0)\\ &\quad - (\text{área entre } z = -0.17 \text{ y } z = 0)\\ &= 0.3413 - 0.0675 = 0.2738 = 27\%\end{aligned}$$

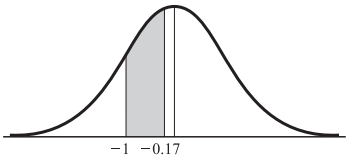


Figura 4-20

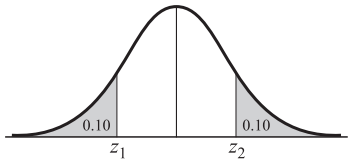


Figura 4-21

b) Sea x_1 la calificación máxima que buscamos y z_1 su equivalente en unidades estándar. De acuerdo con la figura 4-21, el área de la izquierda de z_1 es $10\% = 0.10$; por tanto,

$$\text{Área entre } z_1 \text{ y } 0 = 0.40$$

y $z_1 = -1.28$ (muy cercana).

Entonces, $z_1 = (x_1 - 6.7)/1.2 = -1.28$ y $x_1 = 5.2$ o bien 5 al entero más cercano.

c) Sea x_2 la calificación mínima que se busca y z_2 su equivalente en unidades estándar. De acuerdo con el inciso *b*), por simetría, $z_2 = 1.28$. Por consiguiente $(x_2 - 6.7)/1.2 = 1.28$, y $x_2 = 8.2$ o sea 8 al entero más cercano.

4.59. Un contador Geiger se utiliza para contar partículas radiactivas. Determinar la probabilidad de que en un tiempo t no se cuente ninguna partícula.

La figura 4-22 representa el eje del tiempo con O como origen. La probabilidad de que se cuente una partícula en un tiempo pequeño Δt es proporcional a Δt y, en consecuencia, puede escribirse como $\lambda\Delta t$. Por tanto, la probabilidad de no contar en el tiempo Δt es $1 - \lambda\Delta t$. De manera más precisa, habrá otros términos $(\Delta t)^2$ y órdenes superiores, pero si Δt es pequeño estos términos no se toman en cuenta.

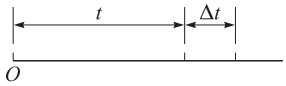


Figura 4-22

Sea $P_0(t)$ la probabilidad de no conteo en el tiempo t. Entonces, $P_0(t + \Delta t)$ es la probabilidad de ningún conteo en el tiempo $t + \Delta t$. Si se supone que las llegadas de las partículas al contador son eventos independientes, la probabilidad de no conteo en el tiempo $t + \Delta t$ es el producto de la probabilidad de no conteo en el tiempo t por la probabilidad de no conteo en el tiempo Δt. Por tanto, sin tomar en cuenta los términos en los que interviene $(\Delta t)^2$ y órdenes superiores, tenemos

$$P_0(t + \Delta t) = P_0(t)[1 - \lambda\Delta t] \tag{1}$$

De (1) obtenemos

$$\lim_{\Delta t \to 0} \frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t) \tag{2}$$

es decir,

$$\frac{dP_0}{dt} = -\lambda P_0 \quad \text{o bien} \quad \frac{dP_0}{P_0} = -\lambda\, dt \tag{3}$$

Resolviendo (3) mediante integración obtenemos

$$\ln P_0 = -\lambda t + c_1 \quad \text{o bien} \quad P_0(t) = ce^{-\lambda t}$$

Para determinar c, observemos que si $t = 0$, $P_0(0) = c$ es la probabilidad de ningún conteo en el tiempo cero, la que evidentemente es 1. Por tanto, $c = 1$ y la probabilidad que buscamos es

$$P_0(t) = e^{-\lambda t} \tag{4}$$

4.60. Volviendo al problema 4.59, calcular la probabilidad de exactamente 1 conteo, contar una partícula, en el tiempo t.

Sea $P_1(t)$ la probabilidad de 1 conteo en el tiempo t. Entonces, $P_1(t + \Delta t)$ es la probabilidad de 1 conteo en el tiempo $t + \Delta t$. Se tendrá un conteo en el tiempo $t + \Delta t$ en los siguientes dos casos mutuamente excluyentes:

i) 1 conteo en el tiempo t y 0 conteos en el tiempo Δt

ii) 0 conteo en el tiempo t y 1 conteo en el tiempo Δt

La probabilidad de *i*) es $P_1(t)(1 - \lambda\Delta t)$.
La probabilidad de *ii*) es $P_0(t)\,\lambda\Delta t$.
Por tanto, sin tomar en cuenta los términos en los que interviene $(\Delta t)^2$ y términos de orden superior,

$$P_1(t + \Delta t) = P_1(t)(1 - \lambda\Delta t) + P_0(t)\lambda\Delta t \tag{1}$$

Esto lo escribimos como

$$\frac{P_1(t + \Delta t) - P_1(t)}{\Delta t} = \lambda P_0(t) - \lambda P_1(t) \tag{2}$$

Tomando el límite cuando $\Delta t \to 0$ y usando la expresión para $P_0(t)$ que se obtuvo en el problema 4.59, esta expresión se convierte en

$$\frac{dP_1}{dt} = \lambda e^{-\lambda t} - \lambda P_1 \tag{3}$$

o bien

$$\frac{dP_1}{dt} + \lambda P_1 = \lambda e^{-\lambda t} \tag{4}$$

Multiplicando por $e^{\lambda t}$, esta expresión la escribimos como

$$\frac{d}{dt}(e^{\lambda t}P_1) = \lambda \tag{5}$$

al integrar obtenemos

$$P_1(t) = \lambda t e^{-\lambda t} + c_2 e^{-\lambda t} \tag{6}$$

Si $t = 0$ y $P_1(0)$ es la probabilidad de un conteo, contar una partícula, en el tiempo 0, la cual es cero. Usando esto en (6), se encuentra $c_2 = 0$. Por tanto,

$$P_1(t) = \lambda t e^{-\lambda t} \tag{7}$$

Si se continúa de esta manera, puede demostrarse que la probabilidad de exactamente n conteos en el tiempo t está dada por

$$P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} \tag{8}$$

que es la distribución de Poisson.

PROBLEMAS SUPLEMENTARIOS

DISTRIBUCIÓN BINOMIAL

4.61. Encuentre la probabilidad de que en 6 lanzamientos de una moneda legal, se obtengan *a*) 0, *b*) 1, *c*) 2, *d*) 3, *e*) 4, *f*) 5, *g*) 6 caras.

4.62. Calcule la probabilidad de obtener *a*) 2 o más caras, *b*) menos de 4 caras, en un lanzamiento de 6 monedas legales.

4.63. Si X denota el número de caras en un solo lanzamiento de cuatro monedas legales, determine *a*) $P(X = 3)$, *b*) $P(X < 2)$, *c*) $P(X \leq 2)$, *d*) $P(1 < X \leq 3)$.

4.64. De 800 familias con 5 hijos cada una, ¿cuántas se espera que tengan *a*) 3 niños, *b*) 5 niñas, *c*) 2 o 3 niños? Suponga que existen las mismas probabilidades para niños que para niñas.

4.65. Calcule la probabilidad de obtener un total de 11 *a*) una vez, *b*) dos veces, en dos lanzamientos de un par de dados no cargados.

4.66. ¿Cuál es la probabilidad de obtener exactamente una vez la suma de 9 en 3 lanzamientos de un par de dados?

4.67. Calcule la probabilidad de acertar por lo menos a 6 de 10 respuestas en un examen de verdadero o falso.

4.68. Un vendedor de seguros vende pólizas a 5 hombres de la misma edad y en buenas condiciones generales de salud. De acuerdo con las tablas actuariales, la probabilidad de que uno de ellos viva 30 años más es $\frac{2}{3}$. Determine la probabilidad de que sobreviva(n) 30 años *a*) los cinco hombres, *b*) por lo menos 3, *c*) sólo 2, *d*) por lo menos uno.

4.69. Calcule *a*) la media, *b*) la desviación estándar, *c*) el coeficiente de sesgo, *d*) el coeficiente de curtosis, de una distribución binomial en la que $p = 0.7$ y $n = 60$. Interprete los resultados.

4.70. Demuestre que si una distribución binomial en la que $n = 100$ es simétrica, su coeficiente de curtosis es 2.9.

4.71. Evalúe *a*) $\sum (x - \mu)^3 f(x)$, *b*) $\sum (x - \mu)^4 f(x)$ de una distribución binomial.

DISTRIBUCIÓN NORMAL

4.72. En un examen de estadística la media fue 78 y la desviación estándar 10. *a*) Determine las puntuaciones estándar de dos estudiantes cuyas calificaciones fueron 93 y 62, respectivamente. *b*) Determine las calificaciones de dos estudiantes cuyas puntuaciones estándar fueron -0.6 y 1.2, respectivamente.

4.73. Encuentre *a*) la media y *b*) la desviación estándar en un examen en el que las calificaciones 70 y 88 correspondieron a las puntuaciones estándar -0.6 y 1.4, respectivamente.

4.74. Encuentre el área bajo la curva normal entre *a*) $z = -1.20$ y $z = 2.40$, *b*) $z = 1.23$ y $z = 1.87$, *c*) $z = -2.35$ y $z = -0.05$.

4.75. Calcule el área bajo la curva normal *a*) a la izquierda de $z = -1.78$, *b*) a la izquierda de $z = 0.56$, *c*) a la derecha de $z = -1.45$, *d*) que corresponde a $z \geq 2.16$, *e*) que corresponde a $-0.80 \leq z \leq 1.53$, *f*) a la izquierda de $z = -2.52$ y a la derecha de $z = 1.83$.

4.76. Si Z está distribuida normalmente con media 0 y varianza 1, determine *a*) $P(Z \geq -1.64)$, *b*) $P(-1.96 \leq Z \leq 1.96)$, *c*) $P(|Z| \geq 1)$.

4.77. Calcule los valores de z tales que *a*) el área de la derecha de z es 0.2266, *b*) el área a la izquierda de z es 0.0314, *c*) el área entre -0.23 y z es 0.5722, *d*) el área entre 1.15 y z es 0.0730, *e*) el área entre $-z$ y z es 0.9000.

4.78. Determine z_1 si $P(Z \geq z_1) = 0.84$, donde z está distribuida normalmente con media 0 y varianza 1.

4.79. Si X está distribuida normalmente con media 5 y desviación estándar 2, encuentre $P(X > 8)$.

4.80. Si las estaturas de 300 estudiantes están distribuidas normalmente con media 68.0 pulgadas y desviación estándar 3.0 pulgadas, ¿cuántos estudiantes tienen una estatura *a*) mayor de 72 pulgadas, *b*) menor o igual a 64 pulgadas, *c*) entre 65 y 71 pulgadas inclusive, *d*) igual a 68 pulgadas? Suponga que las estaturas se han medido a la pulgada más cercana.

4.81. Si los diámetros de bolas de acero están distribuidos normalmente con media de 0.6140 pulgadas y desviación estándar 0.0025 pulgadas, determine el porcentaje de ellas cuyo diámetro *a*) está entre 0.610 y 0.618 pulgadas inclusive, *b*) es mayor a 0.617 pulgadas, *c*) es menor a 0.608 pulgadas, *d*) es igual a 0.615 pulgadas.

4.82. La calificación media en un examen final fue 72 y la desviación estándar fue 9. A los estudiantes cuyas calificaciones pertenece el 10% superior se les dará una *A*. ¿Cuál es la calificación mínima que deberá tener un estudiante para ser acreedor a una *A*?

4.83. Si un conjunto de mediciones está distribuida normalmente, ¿qué porcentaje de las mismas difiere de la media *a*) en más de media desviación estándar, *b*) en menos de tres cuartos de una desviación estándar?

4.84. Si μ es la media y σ la desviación estándar de un conjunto de mediciones distribuidas normalmente, ¿qué porcentaje de las mediciones *a*) están dentro de un rango $\mu \pm 2\sigma$, *b*) están fuera del rango $\mu \pm 1.2\sigma$, *c*) son mayores que $\mu - 1.5\sigma$?

4.85. En el problema 4.84 encuentre la constante *a* tal que el porcentaje de casos *a*) dentro del rango $\mu \pm a\sigma$ sea 75%, *b*) menores a $\mu - a\sigma$ sea 22%.

APROXIMACIÓN NORMAL A LA DISTRIBUCIÓN BINOMIAL

4.86. Determine la probabilidad de que en 200 lanzamientos de una moneda se obtengan *a*) entre 80 y 120, inclusive, caras, *b*) menos de 90 caras, *c*) menos de 85 o más de 115 caras, *d*) exactamente 100 caras.

4.87. Calcule la probabilidad de que un estudiante pueda acertar las respuestas de *a*) 12 o más de 20, *b*) 24 o más de 40 preguntas en un examen de verdadero o falso.

4.88. Diez por ciento de los tornillos producidos con una máquina tiene algún defecto. Encuentre la probabilidad de que en una muestra aleatoria de 400 tornillos producidos con esta máquina, *a*) como máximo 30, *b*) entre 30 y 50, *c*) entre 35 y 45, *d*) 65 o más de los tornillos tengan algún defecto.

4.89. Determine la probabilidad de obtener más de 25 "sietes" en 100 lanzamientos de un par de dados no cargados.

DISTRIBUCIÓN DE POISSON

4.90. Si 3% de las bombillas eléctricas fabricadas por una empresa tienen algún defecto, encuentre la probabilidad de que en una muestra de 100 bombillas, *a*) 0, *b*) 1, *c*) 2, *d*) 3, *e*) 4, *f*) 5 bombillas tengan algún defecto.

4.91. En el problema 4.90, encuentre la probabilidad de que *a*) más de 5, *b*) entre 1 y 3, *c*) 2 o menos bombillas tengan algún defecto.

4.92. Una bolsa contiene una canica roja y siete blancas. Se extrae una de ellas y se observa qué color tiene. Después se reintegra a la bolsa y se mezcla con las otras. Con base en *a*) la distribución binomial, *b*) la aproximación de Poisson a la distribución binomial, encuentre la probabilidad de que en 8 extracciones como la que se menciona, 3 veces se obtenga una canica roja.

4.93. De acuerdo con la Oficina Nacional de Estadística del Departamento de Salud de Estados Unidos, la cantidad promedio de ahogados por accidente por año en ese país, es de 3.0 por cada 100 000 habitantes. Calcule la probabilidad de que en una ciudad de 200 000 habitantes haya *a*) 0, *b*) 2, *c*) 6, *d*) 8, *e*) entre 4 y 8, *f*) menos de 3 ahogados por accidente por año.

4.94. Demuestre que si X_1 y X_2 son variables de Poisson independientes, cuyos parámetros respectivos son λ_1 y λ_2, entonces $X_1 + X_2$ tiene una distribución de Poisson con parámetro $\lambda_1 + \lambda_2$. (*Sugerencia:* Use la función generadora de momentos.) Generalice el resultado a *n* variables.

DISTRIBUCIÓN MULTINOMIAL

4.95. Se lanza 6 veces un dado no cargado. Encuentre la probabilidad de que *a*) se obtenga un "1", dos "2" y tres "3", *b*) cada vez se obtenga un número diferente.

4.96. En una caja hay una gran cantidad de canicas rojas, blancas, azules y amarillas en la proporción 4:3:2:1. Calcule la probabilidad de que en 10 extracciones *a*) 4 sean rojas, 3 sean blancas, 2 sean azules y 1 sea amarilla, *b*) 8 sean rojas y 2 sean amarillas.

4.97. Determine la probabilidad de que en 4 lanzamientos de un dado no cargado no se obtenga ni un "1" ni un "2" ni un "3".

DISTRIBUCIÓN HIPERGEOMÉTRICA

4.98. Una caja contiene 5 canicas rojas y 10 blancas. Si se toman 8 de ellas en forma aleatoria (sin reemplazo), determine la probabilidad de que *a*) 4 sean rojas, *b*) todas sean blancas, *c*) por lo menos 1 sea roja.

4.99. Si de una baraja normal con 52 cartas se extraen 13 cartas en forma aleatoria (sin devolución), encuentre la probabilidad de que *a*) 6 sean cartas de figura, *b*) ninguna sea carta de figura.

4.100. De 60 estudiantes que solicitan inscripción a una universidad, 40 son del este. Si se toman en forma aleatoria 20 estudiantes que solicitan su inscripción, determine la probabilidad de que *a*) 10, *b*) no más de 2, sean del este.

DISTRIBUCIÓN UNIFORME

4.101. Sea X distribuida uniformemente en $-2 \leq x \leq 2$. Calcule *a*) $P(X < 1)$, *b*) $P(|X - 1| \geq \frac{1}{2})$.

4.102. Determine *a*) el tercer, *b*) el cuarto momento alrededor de la media de una distribución uniforme.

4.103. Determine el coeficiente *a*) de sesgo, *b*) de curtosis de una distribución uniforme.

4.104. Si X y Y son independientes y ambas están distribuidas de modo uniforme en el intervalo de 0 a 1, encuentre $P(|X - Y| \geq \frac{1}{2})$.

DISTRIBUCIÓN DE CAUCHY

4.105. Suponga que X está distribuida de acuerdo con la regla de Cauchy según (29), página 114, con $a = 2$. Calcule *a*) $P(X < 2)$, *b*) $P(X^2 \geq 12)$.

4.106. Demuestre que si X_1 y X_2 son independientes y tienen la misma distribución de Cauchy, entonces su media aritmética también tiene esta distribución.

4.107. Sean X_1 y X_2 independientes y distribuidas de manera normal con media 0 y varianza 1. Demuestre que $Y = X_1/X_2$ tiene la distribución de Cauchy.

DISTRIBUCIÓN GAMMA

4.108. Una variable aleatoria X tiene la distribución gamma con $\alpha = 3$ y $\beta = 2$. Encuentre *a*) $P(X \leq 1)$, *b*) $P(1 \leq X \leq 2)$.

DISTRIBUCIÓN JI CUADRADA

4.109. Para una distribución ji cuadrada con 12 grados de libertad, encuentre el valor de χ_c^2 tal que *a*) el área a la derecha de χ_c^2 es 0.05, *b*) el área a la izquierda de χ_c^2 es 0.99, *c*) el área a la derecha de χ_c^2 es 0.025.

4.110. Determine los valores de χ^2 para los que el área de la cola de la derecha de la distribución de χ^2 es 0.05, si el número de grados de libertad ν es igual a *a*) 8, *b*) 19, *c*) 28, *d*) 40.

4.111. Resuelva otra vez el problema 4.110 si el área de la cola de la derecha es 0.01.

4.112. *a*) Calcule χ_1^2 y χ_2^2 tales que el área bajo la distribución χ^2 correspondiente a $\nu = 20$ entre χ_1^2 y χ_2^2 sea 0.95, si se supone que son iguales las áreas a la derecha de χ_2^2 y a la izquierda de χ_1^2, *b*) demuestre que, sin la suposición de áreas iguales en el inciso *a*), los valores de χ_1^2 y χ_2^2 no son únicos.

4.113. Si la variable U tiene distribución ji cuadrada con $\nu = 7$, calcule χ_1^2 y χ_2^2 tales que *a*) $P(U > \chi_2^2) = 0.025$, *b*) $P(U < \chi_1^2) = 0.50$, *c*) $P(\chi_1^2 \leq U \leq \chi_2^2) = 0.90$.

4.114. Determine *a*) $\chi_{0.05}^2$ y *b*) $\chi_{0.95}^2$ para $\nu = 150$.

4.115. Encuentre *a*) $\chi_{0.025}^2$ y *b*) $\chi_{0.975}^2$ para $\nu = 250$.

DISTRIBUCIÓN *t* DE STUDENT

4.116. Dada una distribución t de Student con 15 grados de libertad, calcule el valor de t_1 tal que *a*) el área a la derecha de t_1 sea 0.01, *b*) el área a la izquierda de t_1 sea 0.95, *c*) el área a la derecha de t_1 sea 0.10, *d*) el área a la derecha de t_1 junto con el área a la izquierda de $-t_1$ sea 0.01, *e*) el área entre a $-t_1$ y t_1 sea 0.95.

4.117. Calcule los valores de t para los que el área en la cola derecha de la distribución t sea 0.01, si el número de grados de libertad ν es igual a *a*) 4, *b*) 12, *c*) 25, *d*) 60, *e*) 150.

4.118. Determine los valores de t_1 en la distribución t de Student que satisfacen cada una de las condiciones siguientes: *a*) el área entre $-t_1$ y t_1 sea 0.90 y $\nu = 25$, *b*) el área a la izquierda de $-t_1$ sea 0.025 y $\nu = 20$, *c*) el área a la derecha de t_1 junto con el área a la izquierda de $-t_1$ sea 0.01 y $\nu = 5$, *d*) el área a la derecha de t_1 sea 0.55 y $\nu = 16$.

4.119. Si una variable U tiene la distribución de Student con $\nu = 10$, encuentre la constante c tal que *a*) $P(U > c) = 0.05$, *b*) $P(-c \leq U \leq c) = 0.98$, *c*) $P(U \leq c) = 0.20$, *d*) $P(U \geq c) = 0.90$.

DISTRIBUCIÓN *F*

4.120. Evalúe cada uno de los incisos siguientes:
a) $F_{0.95,15,12}$; *b*) $F_{0.99,120,60}$; *c*) $F_{0.99,60,24}$; *d*) $F_{0.01,30,12}$; *e*) $F_{0.05,9,20}$; *f*) $F_{0.01,8,8}$.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

4.75. *a*) 0.0375 *b*) 0.7123 *c*) 0.9265 *d*) 0.0154 *e*) 0.7251 *f*) 0.0395

4.76. *a*) 0.9495 *b*) 0.9500 *c*) 0.6826

4.77. *a*) 0.75 *b*) -1.86 *c*) 2.08 *d*) 1.625 o 0.849 *e*) ± 1.645

4.78. -0.995 **4.79.** 0.0668

4.80. *a*) 20 *b*) 36 *c*) 227 *d*) 40

4.81. *a*) 93% *b*) 8.1% *c*) 0.47% *d*) 15% **4.82.** 84

4.83. *a*) 61.7% *b*) 54.7% **4.84.** *a*) 95.4% *b*) 23.0% *c*) 93.3%

4.85. *a*) 1.15 *b*) 0.77 **4.86.** *a*) 0.9962 *b*) 0.0687 *c*) 0.0286 *d*) 0.0558

4.87. *a*) 0.2511 *b*) 0.1342 **4.88.** *a*) 0.0567 *b*) 0.9198 *c*) 0.6404 *d*) 0.0079

4.89. 0.0089 **4.90.** *a*) 0.04979 *b*) 0.1494 *c*) 0.2241 *d*) 0.2241 *e*) 0.1680 *f*) 0.1008

4.91. *a*) 0.0838 *b*) 0.5976 *c*) 0.4232 **4.92.** *a*) 0.05610 *b*) 0.06131

4.93. *a*) 0.00248 *b*) 0.04462 *c*) 0.1607 *d*) 0.1033 *e*) 0.6964 *f*) 0.0620

4.95. *a*) 5/3 888 *b*) 5/324 **4.96.** *a*) 0.0348 *b*) 0.000295

4.97. 1/16 **4.98.** *a*) 70/429 *b*) 1/143 *c*) 142/143

4.99. *a*) $\binom{13}{6}\binom{39}{7}\Big/\binom{52}{13}$ *b*) $\binom{13}{0}\binom{39}{13}\Big/\binom{52}{13}$

4.100. *a*) $\binom{40}{10}\binom{20}{10}\Big/\binom{60}{20}$ *b*) $[({}_{40}C_0)({}_{20}C_{20}) + ({}_{40}C_1)({}_{20}C_{19}) + ({}_{40}C_2)({}_{20}C_{18})]/{}_{60}C_{20}$

4.101. *a*) 3/4 *b*) 3/4 **4.102.** *a*) 0 *b*) $(b-a)^4/80$

4.103. *a*) 0 *b*) 9/5 **4.104.** 1/4

4.105. *a*) 3/4 *b*) 1/3 **4.108.** *a*) $1 - \dfrac{13}{8\sqrt{e}}$ *b*) $\dfrac{13}{8}e^{-1/2} - \dfrac{5}{2}e^{-1}$

4.109. *a*) 21.0 *b*) 26.2 *c*) 23.3 **4.110.** *a*) 15.5 *b*) 30.1 *c*) 41.3 *d*) 55.8

4.111. *a*) 20.1 *b*) 36.2 *c*) 48.3 *d*) 63.7 **4.112.** *a*) 9.59 y 34.2

4.113. *a*) 16.0 *b*) 6.35 *c*) áreas iguales asumidas en las dos colas, $\chi_1^2 = 2.17$ y $\chi_2^2 = 14.1$

4.114. *a*) 122.5 *b*) 179.2 **4.115.** *a*) 207.7 *b*) 295.2

4.116. *a*) 2.60 *b*) 1.75 *c*) 1.34 *d*) 2.95 *e*) 2.13

4.117. *a*) 3.75 *b*) 2.68 *c*) 2.48 *d*) 2.39 *e*) 2.33

4.118. *a*) 1.71 *b*) 2.09 *c*) 4.03 *d*) −0.128

4.119. *a*) 1.81 *b*) 2.76 *c*) −0.879 *d*) −1.37

4.120. *a*) 2.62 *b*) 1.73 *c*) 2.40 *d*) 0.352 *e*) 0.340 *f*) 0.166

Parte II

CAPÍTULO 5

POBLACIÓN Y MUESTRA. INFERENCIA ESTADÍSTICA

A menudo, en la práctica interesa obtener conclusiones válidas acerca de un gran grupo de individuos o de objetos. En vez de examinar a todo el grupo, llamado *población*, lo que resulta difícil o imposible, se puede examinar sólo una pequeña parte de esta población, que se llama *muestra*. Esto se hace con el objetivo de deducir ciertos hechos acerca de la población a partir de los resultados encontrados en la muestra, a este proceso se le conoce como *inferencia estadística*. El proceso de obtener muestras se llama *muestreo*.

EJEMPLO 5.1 Se desea obtener conclusiones sobre las estaturas (o los pesos) de 12 000 estudiantes adultos (la población) examinando sólo 100 estudiantes (una muestra) tomada de esta población.

EJEMPLO 5.2 Se desea obtener conclusiones sobre el porcentaje de los pernos defectuosos, que produce una fábrica durante 6 días por semana, examinando 20 pernos cada día producidos en diversas horas. En este caso todos los pernos que se produjeron durante la semana constituyen la población, mientras que los 120 pernos seleccionados constituyen una muestra.

EJEMPLO 5.3 Se desea obtener conclusiones sobre la legalidad de una moneda lanzándola varias veces. La población consta de todos los posibles lanzamientos de la moneda. Una muestra se puede obtener examinando, por ejemplo, los primeros 60 lanzamientos de la moneda y observando los porcentajes de caras y cruces.

EJEMPLO 5.4 Se desea obtener conclusiones sobre los colores de 200 canicas (la población) que contiene una urna. Para ello, se toma una muestra de 20 canicas, la seleccionada se devuelve a la urna después de observar su color.

Hay varios aspectos que es necesario observar. Primero, la palabra *población* no tiene necesariamente el mismo significado que en lenguaje cotidiano, por ejemplo, "la población de cierto lugar es de 180 000 habitantes". Segundo, con frecuencia la palabra *población* suele denotar las observaciones o las medidas y no los individuos o los objetos. En el ejemplo 5.1 se habla de una población de 12 000 estaturas (o pesos), en el ejemplo 5.4 se habla de una población de 200 colores en la urna (algunos de ellos pueden ser iguales). Tercero, la población puede ser finita o infinita, y el número de elementos que la componen se llama *tamaño de la población* y, por lo general, se denota por N. El número de elementos incluidos en la muestra se llama *tamaño de la muestra*, se denota n, y es generalmente finito. En el ejemplo 5.1, $N = 12\,000$, $n = 100$; en el ejemplo 5.3, N es infinita, $n = 60$.

MUESTREO CON Y SIN REEMPLAZO

Si se extrae un objeto de una urna, se tiene la opción de devolverlo o no a ella antes de hacer otra extracción. En el primer caso un objeto particular puede extraerse repetidas veces, mientras que en el segundo sólo puede extraerse una vez. Un muestreo en el que cada miembro de la población puede ser extraído más de una vez se llama *muestreo con reemplazo, con devolución*, mientras que un muestreo en el que cada miembro no puede ser extraído más de una vez se llama *muestreo sin reemplazo.*

Desde un punto de vista teórico, una población finita que se muestrea con reemplazo se considera teóricamente infinita, puesto que pueden extraerse muestras de cualquier tamaño sin agotar la población. Para la mayoría de los propósitos prácticos, el muestreo de una población finita muy grande puede considerarse como un muestreo de una población infinita.

MUESTRAS ALEATORIAS. NÚMEROS ALEATORIOS

Es claro que la confiabilidad de las conclusiones que se obtengan acerca de una población depende de que la muestra se haya elegido de manera adecuada, de tal modo que represente lo mejor posible a la población, y uno de los problemas más importantes de la inferencia estadística es, precisamente, cómo elegir una muestra.

Una manera de hacer esto cuando se trata de poblaciones finitas es asegurarse de que cada miembro de la población tenga la misma posibilidad de pertenecer a la muestra, lo que suele conocerse como *muestreo aleatorio.* El muestreo aleatorio, cuando se trata de poblaciones relativamente pequeñas, puede lograrse sacando números de la suerte o, lo que es equivalente, mediante el uso de una tabla de *números aleatorios* (apéndice H) especialmente elaborada para tal propósito. Vea el problema 5.43.

Como la inferencia de una población no puede ser segura a partir de una muestra, en cualquier enunciado o conclusión debe emplearse el lenguaje de la probabilidad.

PARÁMETROS POBLACIONALES

Una población se considera conocida cuando se sabe cuál es la distribución de probabilidad $f(x)$ (función de probabilidad o función de densidad) de la correspondiente variable aleatoria X. Como, en el ejemplo 5.1, si X es una variable aleatoria cuyos valores son las estaturas (o los pesos) de los 12 000 estudiantes, entonces X tiene la distribución de probabilidad $f(x)$.

Si, por ejemplo, X está distribuida de manera normal, se dice que la población está *distribuida normalmente* o que se tiene una *población normal.* De manera similar, si X tiene distribución binomial, se dice que la población está *distribuida binomialmente* o que se tiene una *población binomial*.

Existen ciertas cantidades que aparecen en $f(x)$, como μ y σ en el caso de la distribución normal y p en el caso de la distribución binomial. Otras cantidades como la mediana, los momentos y el sesgo pueden ser determinadas en términos de éstas. A todas estas cantidades se les suele llamar *parámetros poblacionales.* Cuando se da la población de manera que se conozca $f(x)$, también se conocen los parámetros poblacionales.

Cuando la distribución de probabilidad $f(x)$ no se conoce con precisión, aunque se pueda tener alguna idea de ella o por lo menos se puedan hacer algunas hipótesis en relación con el comportamiento general de $f(x)$, surge un problema importante. Por ejemplo, puede tenerse alguna razón para suponer que una determinada población esté distribuida normalmente. En ese caso, es posible que no se conozca uno o ninguno de los valores μ y σ y que se necesite obtener inferencias estadísticas acerca de ellos.

ESTADÍSTICOS MUESTRALES

Pueden tomarse muestras aleatorias de la población, que después se usen para obtener valores que sirvan para estimar y probar hipótesis acerca de los parámetros poblacionales.

A manera de ilustración, considere el ejemplo 5.1 en donde X es una variable aleatoria cuyos valores son las diversas estaturas. Para obtener una muestra aleatoria de tamaño 100, primero es necesario tomar al azar un individuo de la población. Este individuo puede tener cualquier valor, por ejemplo, x_1 de las diversas estaturas posibles, a x_1 se le llama el valor de la variable aleatoria X_1, donde el subíndice 1 se usa para indicar que corresponde al primer individuo tomado de la población. De manera similar, puede tomarse el segundo individuo para la muestra, el cual puede tener cualquier valor x_2 de los posibles valores de las estaturas y x_2 considerarse como el valor de la variable aleatoria X_2. Este proceso puede continuar hasta X_{100} ya que el tamaño de la muestra es 100. Para simplificar, suponga que el muestreo se hace con reemplazo de manera que un individuo pueda ser tomado más de una vez para integrar la muestra. En este caso, como el tamaño de la muestra es mucho menor que el tamaño de la población, con el muestreo sin reemplazo, devolución, se obtendrán prácticamente los mismos resultados que con el muestreo con reemplazo.

En el caso general, una muestra de tamaño n se describe mediante los valores $x_1, x_2, \ldots, x_n$ de las variables aleatorias $X_1, X_2, \ldots, X_n$. En el caso de un muestreo con reemplazo, $X_1, X_2, \ldots, X_n$ serán variables aleatorias independientes idénticamente distribuidas, con distribución de probabilidad $f(x)$. En consecuencia, su distribución conjunta será

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = f(x_1) f(x_2) \ldots f(x_n) \tag{1}$$

Toda cantidad que se obtiene de una muestra con el propósito de estimar un parámetro poblacional se llama *estadístico muestral* o sólo *estadístico*. Matemáticamente, un estadístico muestral de una muestra de tamaño n puede definirse como una función de las variables aleatorias $X_1, \ldots, X_n$, es decir, $g(X_1, \ldots, X_n)$. La función $g(X_1, \ldots, X_n)$

es otra variable aleatoria, cuyos valores pueden representarse por $g(x_1, \ldots, x_n)$. La palabra *estadístico* suele usarse para la variable aleatoria o para sus valores, pero el significado dependerá del contexto.

En general, a cada parámetro poblacional corresponde un estadístico que se calcula a partir de la muestra. Usualmente, el método para obtener este estadístico a partir de la muestra es similar al método para obtener el parámetro a partir de la población finita, ya que una muestra consta de un conjunto finito de valores. Sin embargo, como se verá, no siempre se obtiene la "mejor estimación", y uno de los problemas importantes de la teoría del muestreo es decidir cómo formar el estadístico muestral adecuado que estime mejor el parámetro poblacional dado. Estos problemas se considerarán en capítulos posteriores.

Siempre que sea posible se usarán letras griegas, como por ejemplo μ y σ, para los valores de los parámetros poblacionales y letras romanas, *m, s*, etc., para los valores correspondientes a los estadísticos muestrales.

DISTRIBUCIONES MUESTRALES

Como se ha visto, un estadístico muestral calculado a partir de $X_1, \ldots, X_n$ es una función de estas variables aleatorias y, por tanto, él mismo es una variable aleatoria. A la distribución de probabilidad de un estadístico muestral suele llamársele *distribución muestral* del estadístico.

Alternativamente, pueden considerarse todas las muestras posibles de tamaño n que pueden obtenerse de una población, y de cada muestra calcular el estadístico. De esta manera se obtiene una distribución del estadístico, que es la distribución muestral.

Para una distribución muestral puede calcularse, por supuesto, la media, la varianza, la desviación estándar, los momentos, etc. A la desviación estándar algunas veces también se le suele llamar *error estándar.*

MEDIA MUESTRAL

Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes, idénticamente distribuidas, obtenidas de una muestra de tamaño n como ya se describió. Entonces, la *media de la muestra* o *media muestral* es una variable aleatoria definida por

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \tag{2}$$

en analogía con (3), página 75. Si $x_1, x_2, \ldots, x_n$ denotan los valores que se obtuvieron de una determinada muestra de tamaño n, entonces la media de esas muestras las representa

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{3}$$

EJEMPLO 5.5 Si en una muestra de tamaño 5 se obtienen los valores muestrales 7, 9, 1, 6 y 2, la media muestral es

$$\bar{x} = \frac{7 + 9 + 1 + 6 + 2}{5} = 5$$

DISTRIBUCIÓN MUESTRAL DE MEDIAS

Sea $f(x)$ la distribución de probabilidad de alguna población dada, de la que se toma una muestra de tamaño n. Entonces, es natural preguntarse por la distribución de probabilidad del estadístico muestral $\bar{X}$, que se conoce como *distribución muestral de la media muestral* o *distribución muestral de medias.* Los teoremas siguientes son importantes en relación con este tema.

Teorema 5-1 La media de la distribución muestral de medias, que se denota como $\mu_{\bar{X}}$, está dada por

$$E(\bar{X}) = \mu_{\bar{X}} = \mu \tag{4}$$

donde μ es la media de la población.

El teorema 5-1 demuestra que el valor esperado de la media muestral es la media poblacional.

Teorema 5-2 Si una población es infinita y el muestreo es aleatorio o si la población es finita y el muestreo se hace con reemplazo, entonces la varianza de la distribución muestral de la media, que se denota $\sigma_{\bar{X}}^2$, está dada por

$$E[(\bar{X} - \mu)^2] = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \tag{5}$$

donde σ^2 es la varianza de la población.

Teorema 5-3 Si el tamaño de la población es N, el muestreo se hace sin reemplazo, sin devolución, y el tamaño de la muestra es $n \leq N$, entonces en vez de (5) se utiliza

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N - n}{N - 1}\right) \tag{6}$$

mientras que $\mu_{\bar{X}}$ sigue siendo la dada por (4).

Observe que cuando $N \to \infty$, (6) se reduce a (5).

Teorema 5-4 Si la población de la que se toman las muestras está distribuida normalmente con media μ y varianza σ^2, entonces la media muestral está distribuida de manera normal, con media μ y varianza σ^2/n.

Teorema 5-5 Suponga que la población de la que se toman las muestras tiene una distribución de probabilidad con media μ y varianza σ^2, que no necesariamente es una distribución normal. Entonces, la variable estandarizada correspondiente a $\bar{X}$, está dada por

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \tag{7}$$

es *asintóticamente normal*, es decir,

$$\lim_{n\to\infty} P(Z \leq z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-u^2/2}\,du \tag{8}$$

El teorema 5-5 es una consecuencia del teorema del límite central de la página 112. Aquí se supone que la población es infinita o que el muestreo se hace con reemplazo. De no ser así, lo anterior es correcto si en (7) se sustituye $\sigma/\sqrt{n}$ por $\sigma_{\bar{X}}$ como se da en (6).

DISTRIBUCIÓN MUESTRAL DE PROPORCIONES

Suponga que se tiene una población infinita distribuida binomialmente, donde p y $q = 1 - p$ son las probabilidades respectivas de que cualquier miembro dado de la población, presente o no cierta propiedad. Por ejemplo, la población puede constar de todos los posibles lanzamientos de una moneda legal, en donde la probabilidad del evento cara es $p = \frac{1}{2}$.

Considere todas las muestras posibles de tamaño n extraídas de esta población, y para cada muestra determine el estadístico proporción P de éxitos. En el caso de una moneda, P puede ser la proporción de caras que se obtiene en n lanzamientos. Así se obtiene una *distribución muestral de proporciones* cuya media μ_p y su desviación estándar σ_p son

$$\mu_p = p \qquad \sigma_p = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1 - p)}{n}} \tag{9}$$

las cuales pueden obtenerse de (4) y (5) utilizando $\mu = p$, $\sigma = \sqrt{pq}$.

En el caso de valores grandes de n ($n \geq 30$), la distribución muestral tiende a una distribución normal, como se ve en el teorema 5-5.

Para poblaciones finitas, cuando el muestreo se hace sin reemplazo, la segunda ecuación de (9) se sustituye por $\sigma_{\bar{x}}$, como se da en (6) por $\sigma = \sqrt{pq}$.

Observe que las ecuaciones en (9) se obtienen más fácilmente si se dividen entre n, la media y la desviación estándar (np y $\sqrt{npq}$) de la distribución binomial.

DISTRIBUCIÓN MUESTRAL DE DIFERENCIAS Y SUMAS

Suponga que se tienen dos poblaciones. En cada muestra de tamaño n_1 obtenida de la primera población se calcula un estadístico S_1; de esta manera se obtiene una distribución muestral de S_1 cuya media y desviación estándar se denotan por μ_{S_1} y σ_{S_1} respectivamente. De manera similar, en cada muestra de tamaño n_2 que se obtuvo de la segunda población se calcula un estadístico S_2 cuya media y desviación estándar son μ_{S_2} y σ_{S_2}, respectivamente.

Tomando todas las combinaciones posibles de estas muestras de las dos poblaciones, se obtiene una distribución de las diferencias, $S_1 - S_2$, a la que se le llama *distribución muestral de diferencias* de los estadísticos. La media y la desviación estándar de esta distribución muestral, que se denotan, respectivamente, $\mu_{S_1-S_2}$ y $\sigma_{S_1-S_2}$, son

$$\mu_{S_1-S_2} = \mu_{S_1} - \mu_{S_2} \qquad \sigma_{S_1-S_2} = \sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2}, \tag{10}$$

siempre que las muestras elegidas no sean de alguna manera dependientes una de la otra, es decir, siempre que las muestras sean *independientes* (en otras palabras, que las variables aleatorias S_1 y S_2 sean independientes).

Si, por ejemplo, S_1 y S_2 son las medias muestrales de las dos poblaciones, que se denotan $\bar{X}_1$ y $\bar{X}_2$, respectivamente, entonces, en el caso de poblaciones infinitas con media y desviación estándar μ_1, σ_1 y μ_2, σ_2, respectivamente, la distribución muestral de las diferencias de medias es

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2, \qquad \sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{11}$$

usando (4) y (5). Este resultado es también válido para poblaciones finitas si el muestreo se hace con reemplazo. La variable estandarizada

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \tag{12}$$

en ese caso es muy próxima a la distribución normal si n_1 y n_2 son grandes (n_1, $n_2 \geq 30$). Resultados similares se obtienen con poblaciones finitas en las que el muestreo se hace sin reemplazo, usando (4) y (6).

Los resultados correspondientes se obtienen para las distribuciones muestrales de las diferencias de proporciones entre dos poblaciones distribuidas de manera binomial con parámetros p_1 , q_1 y p_2 , q_2, respectivamente. En este caso, S_1 y S_2 corresponden a las proporciones de éxito P_1 y P_2, y la ecuación (11) se convierte en

$$\mu_{P_1-P_2} = \mu_{P_1} - \mu_{P_2} = p_1 - p_2, \qquad \sigma_{P_1-P_2} = \sqrt{\sigma_{P_1}^2 + \sigma_{P_2}^2} = \sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} \tag{13}$$

En ocasiones, en lugar de las diferencias entre estadísticos lo que interesa es la suma de los estadísticos. En este caso, la *distribución muestral de la suma de los estadísticos* S_1 y S_2 tiene la media y la desviación estándar dada por

$$\mu_{S_1+S_2} = \mu_{S_1} + \mu_{S_2} \qquad \sigma_{S_1+S_2} = \sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2} \tag{14}$$

si se supone que las muestras son independientes se pueden obtener resultados similares a (11).

VARIANZA MUESTRAL

Si $X_1, X_2, \ldots, X_n$ denotan variables aleatorias de una muestra aleatoria de tamaño n, entonces la variable aleatoria dada por la *varianza de la muestra* o la *varianza muestral* se define de manera análoga a (14), página 77, por

$$S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{n} \tag{15}$$

Ahora, en el teorema 5-1 se encuentra que $E(\bar{X}) = \mu$ y sería deseable que se pudiera tener $E(S^2) = \sigma^2$. Siempre que el valor esperado de un estadístico es igual al parámetro poblacional correspondiente, se dice que el estadístico es un *estimador insesgado* y que su valor es una *estimación insesgada* de este parámetro. Sin embargo, resulta que (vea el problema 5.20)

$$E(S^2) = \mu_{S^2} = \frac{n-1}{n}\sigma^2 \tag{16}$$

el cual es un valor muy cercano a σ^2 sólo con valores grandes de n (por ejemplo, $n \geq 30$). El estimador insesgado deseado se define por

$$\hat{S}^2 = \frac{n}{n-1}S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{n-1} \tag{17}$$

de manera que

$$E(\hat{S}^2) = \sigma^2 \tag{18}$$

Debido a esto algunos especialistas en estadística definen la varianza muestral mediante $\hat{S}^2$ y no mediante S^2 y simplemente sustituyen, en el denominador de (15), n por $n - 1$. Sin embargo, en este texto siempre se definirá la varianza muestral como en (15) ya que, haciéndolo así, muchos de los resultados posteriores se simplifican.

EJEMPLO 5.6 Volviendo al ejemplo 5.5 de la página 155, la varianza muestral es

$$s^2 = \frac{(4-6)^2 + (7-6)^2 + (5-6)^2 + (8-6)^2 + (6-6)^2}{5} = 2$$

mientras que el estimador insesgado es

$$\hat{s}^2 = \frac{5}{4}s^2 = \frac{(4-6)^2 + (7-6)^2 + (5-6)^2 + (8-6)^2 + (6-6)^2}{4} = 2.5$$

El resultado anterior es válido si el muestreo es de una población infinita o de una población finita con reemplazo. Si se muestrea sin reemplazo una población finita de tamaño N, entonces la media de la distribución muestral de las varianzas es

$$E(S^2) = \mu_{S^2} = \left(\frac{N}{N-1}\right)\left(\frac{n-1}{n}\right)\sigma^2 \tag{19}$$

Cuando $N \to \infty$, esto se reduce a (16).

DISTRIBUCIÓN MUESTRAL DE VARIANZAS

Tomando todas las posibles muestras aleatorias de tamaño n de una población y calculando la varianza de cada muestra, se obtiene una distribución muestral de las varianzas. En lugar de hallar la distribución muestral de S^2 o $\hat{S}^2$, es conveniente obtener la distribución muestral de la variable aleatoria

$$\frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{\sigma^2} \tag{20}$$

La distribución de esta variable aleatoria se describe en el teorema siguiente.

Teorema 5-6 Si de una población que tiene una distribución normal se toman muestras aleatorias de tamaño n, entonces la variable muestral (20) tiene una distribución ji cuadrada con $n - 1$ grados de libertad.

Debido al teorema 5-6, la variable en (20) suele denotarse por χ^2. Una prueba de este teorema se encuentra en el problema 5.22.

CASO EN EL QUE NO SE CONOCE LA VARIANZA POBLACIONAL

En los teoremas 5-4 y 5-5 se encuentra que la variable estandarizada

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \tag{21}$$

está distribuida normalmente si la población de la que se toman las muestras de tamaño n también lo está, mientras que si la población no es normal, la variable estandarizada es asintóticamente normal, siempre que $n \geq 30$. En (21) se ha supuesto que se conoce la varianza poblacional σ^2.

Resulta natural preguntar qué ocurre cuando no se conoce la varianza poblacional. Una posibilidad es estimar la varianza poblacional con base en la varianza muestral y colocar después la correspondiente desviación estándar en (21). Una idea mejor es sustituir en (21) la σ por la variable aleatoria $\hat{S}$ que da la desviación estándar muestral y después buscar la distribución del estadístico correspondiente, que aquí se designa por

$$T = \frac{\bar{X} - \mu}{\hat{S}/\sqrt{n}} = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \tag{22}$$

Entonces puede demostrarse, mediante el teorema 4-6 de la página 116, que T tiene la distribución t de Student con $n - 1$ grados de libertad, siempre que la variable aleatoria poblacional esté distribuida de manera normal. Esta conclusión se afirma en el teorema siguiente que se demuestra en el problema 5.24.

Teorema 5-7 Si de una población distribuida normalmente se toman muestras aleatorias de tamaño n, el estadístico (22) tiene la distribución de Student con $n - 1$ grados de libertad.

DISTRIBUCIÓN MUESTRAL DE RAZONES DE VARIANZAS

En la página 157 se indicó cómo obtener distribuciones muestrales de diferencias, en particular de diferencias de medias. Con base en la misma idea, puede obtenerse la distribución muestral de diferencias de varianzas, por ejemplo, $S_1^2 - S_2^2$. Sin embargo, se encuentra que esta distribución muestral es bastante complicada. En lugar de esto, suele considerarse el estadístico S_1^2/S_2^2, ya que una razón grande o pequeña indicará una diferencia grande mientras que una razón casi igual a 1 indica una diferencia pequeña.

Teorema 5-8 Se toman dos muestras aleatorias independientes de tamaños m y n, respectivamente, de dos poblaciones normales cuyas varianzas correspondientes son σ_1^2 y σ_2^2. Entonces, si las varianzas de las muestras aleatorias son S_1^2 y S_2^2, respectivamente, el estadístico

$$F = \frac{mS_1^2/(m-1)\sigma_1^2}{nS_2^2/(n-1)\sigma_2^2} = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \tag{23}$$

tiene la distribución F con $m - 1$ y $n - 1$ grados de libertad.

OTROS ESTADÍSTICOS

Además de la media y la varianza o la desviación estándar, pueden encontrarse muchos otros estadísticos de las muestras. Algunos ejemplos son la mediana, la moda, los momentos, el sesgo y la curtosis. Sus definiciones son análogas a las dadas en el capítulo 3 para las poblaciones. Por lo general pueden hallarse distribuciones muestrales de estos estadísticos, o por lo menos sus medias y desviaciones estándar (errores estándar). En la tabla 5-1 se muestran algunos de éstos.

Tabla 5-1
Errores estándar de algunos estadísticos muestrales

Estadístico muestral	Error estándar	Observaciones
Medias	$\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}$	Esto es válido para muestras grandes o pequeñas en los casos en que la población sea infinita o el muestreo se haga con reemplazo. La distribución muestral de medias es casi normal (asintóticamente normal) para $n \geq 30$, aun cuando la población no sea normal. $\mu_{\bar{X}} = \mu$, la media poblacional en todos los casos.
Proporciones	$\sigma_P = \sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{pq}{n}}$	Las observaciones hechas para las medias también se aplican aquí. $\mu_p = p$ en todos los casos.
Medianas	$\sigma_{\text{med}} = \sigma\sqrt{\dfrac{\pi}{2n}} = \dfrac{1.2533\sigma}{\sqrt{n}}$	Para $n \geq 30$, la distribución muestral de las medianas es casi normal. Los resultados dados son válidos sólo si la población es normal o aproximadamente normal. $\mu_{\text{med}} = \mu$
Desviaciones estándar	(1) $\sigma_S = \dfrac{\sigma}{\sqrt{2n}}$ (2) $\sigma_S = \sqrt{\dfrac{\mu_4 - \sigma^4}{4n\sigma^2}}$	Para $n \geq 100$, la distribución muestral de S se asemeja mucho a la normal. σ_S es la dada por (1) sólo si la población es normal (o aproximadamente normal). Si la población no es normal, puede usarse (2). Observe que (2) se reduce a (1) cuando $\mu_4 = 3\sigma^4$, lo cual es válido para poblaciones normales. Para $n \geq 100$, $\mu_S = \sigma$, aproximadamente.
Varianzas	(1) $\sigma_{S^2} = \sigma^2\sqrt{\dfrac{2}{n}}$ (2) $\sigma_{S^2} = \sqrt{\dfrac{\mu_4 - \sigma^2}{n}}$	Las observaciones hechas para las desviaciones estándar también se aplican aquí. Observe que (2) produce (1) en caso de que la población sea normal. $\mu_{S^2} = (n-1)\sigma^2/n$ lo que para muestras grandes de $n(n \geq 30)$ da un valor muy cercano a σ^2.

DISTRIBUCIONES DE FRECUENCIAS

Si una muestra (o una población) es grande, es difícil observar sus diversas características o calcular estadísticos como la media o la desviación estándar. Por esta razón es útil organizar o agrupar los *datos en bruto*. Para ilustrar esto, suponga que se tiene una muestra que consta de las estaturas de 100 estudiantes varones de la universidad XYZ. Estos datos se organizan en *clases* o *categorías* y se determina la cantidad de individuos que pertenece a cada clase, lo que se llama la *frecuencia de clase.* El resultado, que se muestra en la tabla 5-2, se conoce como *distribución de frecuencias* o *tabla de frecuencias.*

La primera clase o categoría, por ejemplo, consta de las estaturas que van de 60 a 62 pulgadas, lo que se anota como 60-62, y se le conoce como *intervalo de clase.* Como hay 5 estudiantes cuya estatura pertenece a esta clase, la frecuencia de clase correspondiente es 5. Puesto que una estatura registrada de 60 pulgadas se encuentra en realidad entre 59.5 y 60.5 pulgadas, mientras que una estatura de 62 pulgadas está en realidad entre 61.5 y 62.5 pulgadas, este intervalo de clase podría haber sido también 59.5-62.5. El siguiente intervalo de clase habría sido entonces 62.5-65.5, etc. En el intervalo de clase 59.5-62.5, a los números 59.5 y 62.5 se les llama *límites de clase*. La amplitud del intervalo de clase j-ésimo, que se denota con c_j, que por lo general es el mismo para todas las clases (en cuyo caso se denota c), es la diferencia entre el límite de clase superior y el límite inferior. En este caso, $c = 62.5 - 59.5 = 3$.

Tabla 5-2

Estaturas de 100 estudiantes varones de la universidad XYZ

Estatura (pulgadas)	Número de estudiantes
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
TOTAL	100

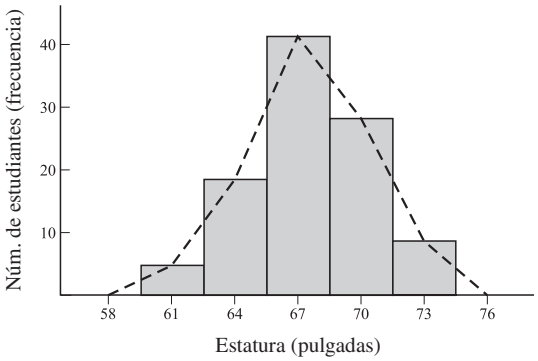


Figura 5-1

El punto medio del intervalo de clase, que se toma como representativo de la clase, se llama *marca de clase.* En la tabla 5-2, la marca de clase correspondiente al intervalo de clase 60-62 es 61.

Una gráfica de la distribución de frecuencias puede ser un histograma, como el que se muestra sombreado en la figura 5-1, o un *polígono de frecuencias* en el que se unen los puntos medios de las partes superiores de los rectángulos del histograma. Es interesante observar que la forma de la gráfica parece indicar que la muestra se ha obtenido de una población de estaturas distribuida normalmente.

DISTRIBUCIÓN DE FRECUENCIAS RELATIVAS

En la tabla 5-2 se dividió el número de estudiantes entre el total (100) y se anotaron las frecuencias relativas o porcentuales y no la cantidad de estudiantes en cada clase, el resultado es una *distribución de frecuencias relativas* o *porcentuales*. Por ejemplo, la frecuencia relativa o porcentual de la clase 63-65 es 18/100, es decir 18%. El histograma correspondiente es similar al de la figura 5-1, excepto que el eje vertical corresponde a frecuencias relativas en lugar de frecuencias. La suma de las áreas de los rectángulos es 1 o 100%.

Una distribución de frecuencias relativas puede considerarse como una distribución de probabilidades en la que las probabilidades se sustituyen por frecuencias relativas. Dado que éstas pueden verse como *probabilidades empíricas* (vea la página 5), las distribuciones de frecuencias relativas se conocen como *distribuciones de probabilidad empírica.*

CÁLCULO DE LA MEDIA, VARIANZA Y MOMENTOS PARA DATOS AGRUPADOS

Una distribución de frecuencias puede representarse como en la tabla 5-3, en la cual se proporciona cada marca de clase y la frecuencia de clase correspondiente. La frecuencia total es n, es decir,

$$n = f_1 + f_2 + \cdots + f_k = \sum f$$

Tabla 5-3

Marca de clase	Frecuencias de clase
x_1	f_1
x_2	f_2
$\vdots$	$\vdots$
x_k	f_k
TOTAL	n

Dado que hay f_1 números iguales a x_1, f_2 números iguales a x_2, . . . , f_k números iguales a x_k, la media es

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \cdots + f_kx_k}{n} = \frac{\sum fx}{n} \tag{24}$$

De manera similar, la varianza es

$$s^2 = \frac{f_1(x_1 - \bar{x})^2 + f_2(x_2 - \bar{x})^2 + \cdots + f_k(x_k - \bar{x})^2}{n} = \frac{\sum f(x - \bar{x})^2}{n} \tag{25}$$

Observe la analogía de (24) y (25) con los resultados (2) de la página 75 y (13) de la página 77, si f_j/n corresponde a probabilidades empíricas.

Para el caso en el que todos los intervalos de clase son del mismo tamaño c, existen métodos más cortos para calcular la media y la varianza. Estos métodos, se conocen como *métodos de compilación*, utilizan la transformación de la marca de clase x en un entero correspondiente u dado por

$$x = a + cu \tag{26}$$

donde a es una marca de clase elegida de manera arbitraria que corresponde a $u = 0$. Las fórmulas de compilación para la media y varianza son

$$\bar{x} = a + \frac{c}{n}\sum fu = a + c\bar{u} \tag{27}$$

$$s^2 = c^2\left[\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2\right] = c^2(\bar{u}^2 - \bar{u}^2 = 0) \tag{28}$$

Para los momentos de orden superior existen fórmulas similares. Los momentos r-ésimos respecto a la media y al origen son, respectivamente

$$m_r = \frac{f_1(x_1 - \bar{x})^r + \cdots + f_k(x_k - \bar{x})^r}{n} = \frac{\sum f(x - \bar{x})^r}{n} \tag{29}$$

$$m'_r = \frac{f_1x_1^r + \cdots + f_kx_k^r}{n} = \frac{\sum fx^r}{n} \tag{30}$$

Estos dos tipos de momentos se relacionan mediante

$$\begin{aligned} m_1 &= 0 \\ m_2 &= m'_2 - m'^2_1 \\ m_3 &= m'_3 - 3m'_1m'_2 + 2m'^3_1 \\ m_4 &= m'_4 - 4m'_1m'_3 + 6m'^2_1m'_2 - 3m'^4_1 \end{aligned} \tag{31}$$

etc. Si se escribe

$$M_r = \frac{\sum f(u - \bar{u})^r}{n} \qquad M'_r = \frac{\sum fu^r}{n}$$

donde u está dada por (26), las relaciones (31) también son válidas con las M. Pero

$$m_r = \frac{\sum f(x - \bar{x})^r}{n} = \frac{\sum f[(a + cu) - (a + c\bar{u})]^r}{n} = \frac{\sum fc^r(u - \bar{u})^r}{n} = c^rM_r$$

de manera que de (31) se obtienen las fórmulas de compilación

$$\begin{aligned} m_1 &= 0 \\ m_2 &= c^2(M'_2 - M'^2_1) \\ m_3 &= c^3(M'_3 - 3M'_1M'_2 + 2M'^3_1) \\ m_4 &= c^4(M'_4 - 4M'_1M'_3 + 6M'^2_1M'_2 - 3M'^4_1) \end{aligned} \tag{32}$$

etc. Por supuesto que la segunda ecuación de (32) es la misma que (28).

De manera similar, para muestras agrupadas, pueden encontrarse otros estadísticos como el sesgo y la curtosis.

PROBLEMAS RESUELTOS

DISTRIBUCIÓN MUESTRAL DE MEDIAS

5.1. Una población consta de cinco números: 2, 3, 6, 8, 11. Considerar todas las muestras posibles de tamaño dos que puedan extraerse con reemplazo de esta población. Encontrar *a*) la media de la población, *b*) la desviación estándar de la población, *c*) la media de la distribución muestral de medias, *d*) la desviación estándar de la distribución muestral de medias, es decir, el error estándar de las medias.

a)
$$\mu = \frac{2 + 3 + 6 + 8 + 11}{5} = \frac{30}{5} = 6.0$$

b)
$$\sigma^2 = \frac{(2 - 6)^2 + (3 - 6)^2 + (6 - 6)^2 + (8 - 6)^2 + (11 - 6)^2}{5} = \frac{16 + 9 + 0 + 4 + 25}{5} = 10.8$$

y $\sigma = 3.29$.

c) Hay $5(5) = 25$ muestras de tamaño dos que pueden extraerse con reemplazo (ya que cada uno de los cinco números de la primera extracción puede combinarse con cualquiera de los cinco números de la segunda extracción). Estas muestras son

(2, 2)	(2, 3)	(2, 6)	(2, 8)	(2, 11)
(3, 2)	(3, 3)	(3, 6)	(3, 8)	(3, 11)
(6, 2)	(6, 3)	(6, 6)	(6, 8)	(6, 11)
(8, 2)	(8, 3)	(8, 6)	(8, 8)	(8, 11)
(11, 2)	(11, 3)	(11, 6)	(11, 8)	(11, 11)

Las correspondientes medias muestrales son

(1)

2.0	2.5	4.0	5.0	6.5
2.5	3.0	4.5	5.5	7.0
4.0	4.5	6.0	7.0	8.5
5.0	5.5	7.0	8.0	9.5
6.5	7.0	8.5	9.5	11.0

y la media de la distribución muestral de medias es

$$\mu_{\bar{X}} = \frac{\text{suma de todas las medias muestrales en (1)}}{25} = \frac{150}{25} = 6.0$$

lo que ilustra el hecho de que $\mu_{\bar{X}} = \mu$. Para una prueba general de esto vea el problema 5.6.

d) La varianza $\sigma_{\bar{X}}^2$ de la distribución muestral de medias se obtiene restando la media 6 de cada uno de los números dados en (1), elevando al cuadrado el resultado, se suman los 25 números que se obtuvieron y se divide entre 25. El resultado final es

$$\sigma_{\bar{X}}^2 = \frac{135}{25} = 5.40 \text{ de manera que } \sigma_{\bar{X}} = \sqrt{5.40} = 2.32$$

Esto ilustra el hecho de que en una población finita en la que se muestree con reemplazo (o en una población infinita), $\sigma_{\bar{X}}^2 = \sigma^2/n$, ya que el lado derecho es $10.8/2 = 5.40$, lo que coincide con el valor que ya encontramos. Para una prueba general de esto, vea el problema 5.7.

5.2. Resolver el problema 5.1 en el caso que el muestreo se haga sin reemplazo.

Como en *a*) y *b*) del problema 5.1, $\mu = 6$ y $\sigma^2 = 10.8$ $\sigma = 3.29$.

c) Hay ${}_5C_2 = 10$ muestras de tamaño dos que pueden obtenerse sin reemplazo (esto significa que se saca un número y después otro número diferente al primero) de esta población, y son:

(2, 3), (2, 6), (2, 8), (2, 11), (3, 6), (3, 8), (3, 11), (6, 8), (6, 11), (8, 11).

La muestra (2, 3), por ejemplo, se considera la misma que la muestra (3, 2).

Las medias muestrales correspondientes son

$$2.5, \quad 4.0, \quad 5.0, \quad 6.5, \quad 4.5, \quad 5.5, \quad 7.0, \quad 7.0, \quad 8.5, \quad 9.5$$

y la media de la distribución muestral de medias es

$$\mu_{\bar{X}} = \frac{2.5 + 4.0 + 5.0 + 6.5 + 4.5 + 5.5 + 7.0 + 7.0 + 8.5 + 9.5}{10} = 6.0$$

lo que ilustra el hecho de que $\mu_{\bar{X}} = \mu$.

d) La varianza de la distribución muestral de medias es

$$\sigma_{\bar{X}}^2 = \frac{(2.5 - 6.0)^2 + (4.0 - 6.0)^2 + (5.0 - 6.0)^2 + \cdots + (9.5 - 6.0)^2}{10} = 4.05$$

y $\sigma_{\bar{X}} = 2.01$.

Esto ilustra que $\sigma_{\bar{X}}^2 = \dfrac{\sigma^2}{n}\left(\dfrac{N-n}{N-1}\right)$, ya que el lado derecho es igual a $\dfrac{10.8}{2}\left(\dfrac{5-2}{5-1}\right) = 4.05$, que es lo que obtuvimos antes.

Vea el problema 5.47 para una prueba general de este resultado.

5.3. Se tienen las estaturas de 3 000 estudiantes varones de una universidad, están distribuidas normalmente, con media de 68.0 pulgadas y desviación estándar de 3.0 pulgadas. Si se toman 80 muestras cada una de 25 estudiantes, ¿cuál será la media y desviación estándar de las medias que se obtendrán si se muestrea *a*) con reemplazo, *b*) sin reemplazo?

La cantidad de muestras de tamaño 25 que teóricamente pueden obtenerse de un grupo de 3 000 estudiantes con y sin sustitución son $(3\,000)^{25}$ y ${}_{3000}C_{25}$, respectivamente, lo que es mucho más grande que 80. De manera que no se obtiene una verdadera distribución muestral de medias, sino sólo una distribución muestral *experimental*. De cualquier manera, dado que la cantidad de muestras es grande, las dos distribuciones muestrales serán bastante similares. Por tanto, la media y la desviación estándar de las 80 medias muestrales serán similares a las de la distribución teórica. Por tanto, tenemos

a) $\mu_{\bar{X}} = \mu = 68.0$ pulgadas y $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{3}{\sqrt{25}} = 0.6$ pulgadas

b) $\mu_{\bar{X}} = \mu = 68.0$ pulgadas y $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}} = \dfrac{3}{\sqrt{25}}\sqrt{\dfrac{3\,000 - 25}{3\,000 - 1}}$

lo que sólo es poco menor a 0.6 pulgadas y para propósitos prácticos puede considerarse igual que en el muestreo con reemplazo.

De manera que, puede esperarse que la distribución muestral experimental de medias se distribuya de manera casi normal, y que su media sea 68.0 pulgadas y desviación estándar 0.6 pulgadas.

5.4. ¿En cuántas muestras del problema 5.3 puede esperarse que la media *a*) esté entre 66.8 y 68.3 pulgadas, *b*) sea menor a 66.4 pulgadas?

La media $\bar{X}$ de una muestra en unidades estándar está dada por $Z = \dfrac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \dfrac{\bar{X} - 68.0}{0.6}$.

a) 66.8 en unidades estándar $= (66.8 - 68.0)/0.6 = -2.0$

68.3 en unidades estándar $= (68.3 - 68.0)/0.6 = 0.5$

Proporción de muestras cuyas medias están entre 66.8 y 68.3 pulgadas
$=$ (área bajo la curva normal entre $z = -2.0$ y $z = 0.5$)
$=$ (área entre $z = -2$ y $z = 0$)
$+$ (área entre $z = 0$ y $z = 0.5$)
$= 0.4772 + 0.1915 = 0.6687$

Entonces, la cantidad esperada de muestras $= (80)(0.6687)$ o 53 (figura 5-2).

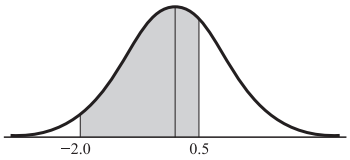


Figura 5-2

b) 66.4 en unidades estándar $= (66.4 - 68.0/0.6 = -2.67)$

Proporción de muestras cuyas medias son menores a 66.4 pulgadas
$=$ (área bajo la curva normal a la izquierda de $z = -2.67$)
$=$ (área a la izquierda de $z = 0$)
$-$ (área entre $z = -2.67$ y $z = 0$)
$= 0.5 - 0.4962 = 0.0038$

Entonces, la cantidad esperada de muestras $= (80)(0.0038) = 0.304$ o cero (figura 5-3).

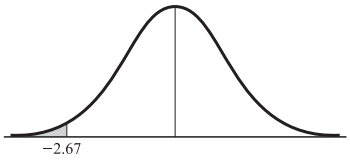


Figura 5-3

5.5. Quinientos balines tienen un peso medio de 5.02 onzas y una desviación estándar de 0.30 onzas. Encontrar la probabilidad de que una muestra aleatoria de 100 balines tomada de este grupo tenga un peso combinado *a*) entre 496 y 500 onzas, *b*) más de 510 onzas.

En la distribución muestral de las medias, $\mu_{\bar{X}} = \mu = 5.02$ onzas.

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{0.30}{\sqrt{100}}\sqrt{\frac{500-100}{500-1}} = 0.027$$

a) El peso combinado de los 100 balines estará entre 496 y 500 onzas si el peso medio de los 100 balines está entre 4.96 y 5.00 onzas (figura 5-4).

$$4.96 \text{ en unidades estándar} = \frac{4.96 - 5.02}{0.027} = -2.22$$

$$5.00 \text{ en unidades estándar} = \frac{5.00 - 5.02}{0.027} = -0.74$$

Probabilidad que buscamos
$=$ (área entre $z = -2.22$ y $z = -0.74$)
$=$ (área entre $z = -2.22$ y $z = 0$)
$-$ (área entre $z = -0.74$ y $z = 0$)
$= 0.4868 - 0.2704 = 0.2164$

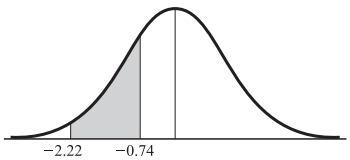


Figura 5-4

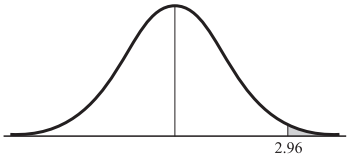


Figura 5-5

b) El peso combinado que excederá 510 onzas si el peso medio de los 100 balines excede 5.10 onzas (figura 5-5).

$$5.10 \text{ en unidades estándar} = \frac{5.10 - 5.02}{0.027} = 2.96$$

Probabilidad que buscamos
= (área a la derecha de $z = 2.96$)
= (área a la derecha de $z = 0$)
− (área entre $z = 0$ y $z = 2.96$)
$= 0.5 - 0.4985 = 0.0015$

Por tanto, sólo hay 3 posibilidades en 2 000 de extraer una muestra de 100 balines con un peso combinado que exceda las 510 onzas.

5.6. Demostrar el teorema 5-1, página 155.

Dado que $X_1, X_2, \ldots, X_n$ son variables aleatorias que tienen la misma distribución que la población, la cual tiene media μ, se cumple

$$E(X_k) = \mu \qquad k = 1, 2, \ldots, n$$

Entonces, dado que la media muestral se define como

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n}$$

se tiene, como se deseaba

$$E(\bar{X}) = \frac{1}{n}[E(X_1) + \cdots + E(X_n)] = \frac{1}{n}(n \cdot \mu) = \mu$$

5.7. Demostrar el teorema 5-2, página 156.

Tenemos

$$\bar{X} = \frac{X_1}{n} + \frac{X_2}{n} + \cdots + \frac{X_n}{n}$$

Entonces, dado que $X_1, \ldots, X_n$ son variables independientes cuya varianza es σ^2, de acuerdo con los teoremas 3-5 y 3-7:

$$\text{Var}(\bar{X}) = \frac{1}{n^2}\text{Var}(X_1) + \cdots + \frac{1}{n^2}\text{Var}(X_n) = n\left(\frac{1}{n^2}\sigma^2\right) = \frac{\sigma^2}{n}$$

DISTRIBUCIÓN MUESTRAL DE PROPORCIONES

5.8. Calcular la probabilidad de que en 120 lanzamientos de una moneda legal *a*) entre 40 y 60% sean caras, o bien, *b*) $\frac{5}{8}$ o más sean caras.

Los 120 lanzamientos de la moneda se consideran como una muestra de la población infinita de todos los posibles lanzamientos de la moneda. En esta población, la probabilidad de cara es $p = \frac{1}{2}$ y la probabilidad de cruz es $q = 1 - p = \frac{1}{2}$.

a) Queremos determinar la probabilidad de que el número de caras en 120 lanzamientos esté entre 40% de 120, es decir 48, y 60% de los 120, es decir 72. Procedemos como en el capítulo 4, aplicamos la aproximación normal a la distribución binomial. Como el número de caras es una variable discreta, la probabilidad que buscamos es que el número de caras se encuentre entre 47.5 y 72.5. (Vea la figura 5-6.)

$$\mu = \text{número esperado de caras} = np = 120\left(\frac{1}{2}\right) = 60$$

y

$$\sigma = \sqrt{npq} = \sqrt{(120)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 5.48$$

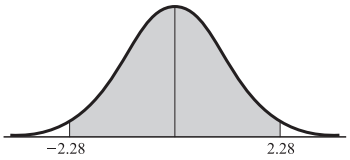


Figura 5-6

$$47.5 \text{ en unidades estándar } = \frac{47.5 - 60}{5.48} = -2.28$$

$$72.5 \text{ en unidades estándar } = \frac{72.5 - 60}{5.48} = 2.28$$

Probabilidad que buscamos
= (área bajo la curva normal
entre $z = -2.28$ y $z = 2.28$)
= 2(área entre $z = 0$ y $z = 2.28$)
= 2(0.4887) = 0.9774

Otro método

$$\mu_P = p = \frac{1}{2} = 0.50 \qquad \sigma_P = \sqrt{\frac{pq}{n}} = \sqrt{\frac{\frac{1}{2}\left(\frac{1}{2}\right)}{120}} = 0.0456$$

$$40\% \text{ en unidades estándar } = \frac{0.40 - 0.50}{0.0456} = -2.19$$

$$60\% \text{ en unidades estándar } = \frac{0.60 - 0.50}{0.0456} = 2.19$$

Por tanto, la probabilidad que buscamos es el área bajo la curva normal entre $z = -2.19$ y $z = 2.19$, es decir, 2(0.4857) = 0.9714.

Aun cuando este resultado es exacto a dos cifras significativas, no coincide con exactitud con el resultado anterior debido a que no se ha usado el hecho de que la proporción es, en realidad, una variable discreta. Para tomar en cuenta esto, se resta $\frac{1}{2n} = \frac{1}{2(120)}$ de 0.40 y se suma $\frac{1}{2n} = \frac{1}{2(120)}$ a 0.60. Por tanto, las proporciones que se buscan en unidades estándar son, dado que $1/240 = 0.00417$,

$$\frac{0.40 - 0.00417 - 0.50}{0.0456} = -2.28 \qquad \text{y} \qquad \frac{0.60 + 0.00417 - 0.50}{0.0456} = 2.28$$

con lo cual se obtiene la coincidencia con el primer método.

Observamos que (0.40 – 0.00417) y (0.60 + 0.00417) corresponden a las proporciones 47.5/120 y 72.5/120 en el primer método anterior.

b) Usando el segundo método del inciso *a*), encontramos que como $\frac{5}{8} = 0.6250$,

$$(0.6250 - 0.00417) \text{ en unidades estándar } = \frac{0.6250 - 0.00417 - 0.50}{0.0456} = 2.65$$

Probabilidad que buscamos = (área bajo la curva normal a la derecha de $z = 2.65$)
= (área a la derecha de $z = 0$)
− (área entre $z = 0$ y $z = 2.65$)
= 0.5 − 0.4960 = 0.0040

5.9. Cada una de las personas de un grupo de 500 lanza una moneda legal 120 veces. ¿Cuántas personas se espera que en sus lanzamientos obtengan entre 40 y 60% de caras, *b*) $\frac{5}{8}$ de sus lanzamientos muestren caras?

Este problema está estrechamente relacionado con el problema 5.8. Aquí se consideran 500 muestras, cada una de tamaño 120, de la población infinita de todos los posibles lanzamientos de una moneda.

a) En el inciso *a*) del problema 5.8 se indica que de todas las muestras posibles, de 120 lanzamientos cada una, puede esperarse que en el 97.74% de ellos, el porcentaje de caras esté entre 40 y 60%. En 500 muestras puede esperarse que alrededor de 97.74% de las 500, es decir 489, muestras tengan esta propiedad. Se deduce que aproximadamente 489 personas obtengan, como resultado del experimento entre 40 y 60% de caras.

Es interesante observar que se espere que 500 – 489 = 11 personas obtengan un porcentaje de caras que no esté entre 40 y 60%. Estas personas pueden concluir, lo que sería razonable, que sus monedas estaban cargadas. Este tipo de error es un *riesgo* que siempre está presente cuando se trata con la probabilidad.

b) Con un razonamiento similar al del inciso *a*), concluimos que, aproximadamente, (500)(0.0040) = 2 personas obtendrán en $\frac{5}{8}$ o más de sus lanzamientos una cara.

5.10. Se ha encontrado que 2% de las herramientas que produce cierta máquina tienen algún defecto. ¿Cuál es la probabilidad de que en 400 de estas herramientas, *a*) 3% o más, *b*) 2% o menos tengan algún defecto?

$$\mu_P = p = 0.02 \qquad \text{y} \qquad \sigma_P = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.02(0.98)}{400}} = \frac{0.14}{20} = 0.007$$

a) Usando la corrección para variables discretas, $1/2n = 1/800 = 0.00125$, se tiene

$$(0.03 - 0.00125) \text{ en unidades estándar } = \frac{0.03 - 0.00125 - 0.02}{0.007} = 1.25$$

$$\text{Probabilidad que buscamos } = (\text{área bajo la curva normal a la derecha de } z = 1.25) = 0.1056$$

Si no se usa la corrección obtenemos 0.0764.

Otro método

(3% de 400) = 12 herramientas con defecto. Considerando valores continuos, 12 o más herramientas significa 11.5 o más.

$$\mu = (2\% \text{ de } 400) = 8 \qquad \sigma = \sqrt{npq} = \sqrt{(400)(0.02)(0.98)} = 2.8$$

Entonces, 11.5 en unidades estándar = (11.5 – 8)/2.8 = 1.25, y como antes, la probabilidad que buscamos es 0.1056.

b) $$(0.02 + 0.00125) \text{ en unidades estándar } = \frac{0.02 + 0.00125 - 0.02}{0.007} = 0.18$$

$$\text{Probabilidad que se busca } = (\text{área bajo la curva normal a la izquierda de } z = 0.18)$$
$$= 0.5000 + 0.0714 = 0.5714$$

Si no se usa la corrección se obtiene 0.5000. También puede usarse el segundo método del inciso *a*).

5.11. En una elección, se encuentra que determinado candidato obtuvo 46% de los votos. Determinar la probabilidad de que en una encuesta de *a*) 200, *b*) 1 000 personas elegidas al azar de la población de los votantes se encuentre una mayoría de votos a favor de este candidato.

a) $$\mu_P = p = 0.46 \qquad \text{y} \qquad \sigma_P = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.46(0.54)}{200}} = 0.0352$$

Como $1/2n = 1/400 = 0.0025$, se tiene una mayoría en la muestra si la proporción a favor del candidato es 0.50 + 0.0025 = 0.5025 o más. (Esta proporción se obtiene también si se considera que 101 o más corresponde a la mayoría, pero esto como variable continua es 100.5; y, de esta manera, la proporción es 100.5/200 = 0.5025.)

Entonces, 0.5025 en unidades estándar = (0.5025 – 0.46)/0.0352 = 1.21 y

$$\text{Probabilidad que buscamos} = (\text{área bajo la curva normal a la derecha de } z = 1.21)$$
$$= 0.5000 - 0.3869 = 0.1131$$

b) $\mu_P = p = 0.46$, $\sigma_P = \sqrt{pq/n} = \sqrt{0.46(0.54)1\,000} = 0.0158$, y

$$0.5025 \text{ en unidades estándar } = \frac{0.5025 - 0.46}{0.0158} = 2.69$$

$$\begin{aligned}\text{Probabilidad que buscamos } &= (\text{área bajo la curva normal a la derecha de } z = 2.69)\\ &= 0.5000 - 0.4964 = 0.0036\end{aligned}$$

DISTRIBUCIÓN MUESTRAL DE DIFERENCIAS Y SUMAS

5.12. Sea U_1 una variable que represente cualquiera de los elementos de la población 3, 7, 8, y U_2 una variable que represente cualquiera de los elementos de la población 2, 4. Calcular *a*) μ_{U_1}, *b*) μ_{U_2}, *c*) $\mu_{U_1-U_2}$, *d*) σ_{U_1}, *e*) σ_{U_2}, *f*) $\sigma_{U_1-U_2}$.

a) $\mu_{U_1} = \text{media de la población } U_1 = \frac{1}{3}(3 + 7 + 8) = 6$

b) $\mu_{U_2} = \text{media de la población } U_2 = \frac{1}{2}(2 + 4) = 3$

c) La población que consta de todas las diferencias entre los miembros de U_1 y los miembros de U_2 es

$$\begin{matrix} 3-2 & 7-2 & 8-2 \\ 3-4 & 7-4 & 8-4 \end{matrix} \quad \text{o} \quad \begin{matrix} 1 & 5 & 6 \\ -1 & 3 & 4 \end{matrix}$$

Entonces, $\mu_{U_1-U_2} = \text{media de } (U_1 - U_2) = \dfrac{1 + 5 + 6 + (-1) + 3 + 4}{6} = 3$

lo que ilustra el resultado general $\mu_{U_1-U_2} = \mu_{U_1} - \mu_{U_2}$, como puede verse en *a*) y *b*).

d) $\sigma^2_{U_1} = \text{varianza de la población } U_1 = \dfrac{(3-6)^2 + (7-6)^2 + (8-6)^2}{3} = \dfrac{14}{3}$

o $\sigma_{U_1} = \sqrt{\dfrac{14}{3}}$.

e) $\sigma^2_{U_2} = \text{varianza de la población } U_2 = \dfrac{(2-3)^2 + (4-3)^2}{2} = 1$

o $\sigma_{U_2} = 1$.

f) $\sigma^2_{U_1-U_2} = \text{varianza de la población } (U_1 - U_2)$

$$= \frac{(1-3)^2 + (5-3)^2 + (6-3)^2 + (-1-3)^2 + (3-3)^2 + (4-3)^2}{6} = \frac{17}{3}$$

o $\sigma_{U_1-U_2} = \sqrt{\dfrac{17}{3}}$.

Esto ilustra el resultado general para muestras independientes, $\sigma_{U_1-U_2} = \sqrt{\sigma^2_{U_1} + \sigma^2_{U_2}}$, como vemos en *d*) y *e*). La demostración del resultado general es una consecuencia del teorema 3-7, página 78.

5.13. La vida media de los bulbos eléctricos producidos por el fabricante *A* es de 1 400 horas y la desviación estándar es 200 horas; la vida media de los bulbos eléctricos producidos por el fabricante *B* es de 1 200 horas y desviación estándar de 100 horas. Si se prueban muestras aleatorias de 125 bulbos de cada marca, ¿cuál es la probabilidad de que los bulbos de la marca *A* tengan una vida media que sea, por lo menos, de *a*) 160 horas, *b*) 250 horas más que las de los bulbos de la marca *B*?

Sean $\bar{X}_A$ y $\bar{X}_B$ las vidas medias de las muestras A y B, respectivamente. Entonces,

$$\mu_{\bar{X}_A-\bar{X}_B} = \mu_{\bar{X}_A} - \mu_{\bar{X}_B} = 1\,400 - 1\,200 = 200 \text{ horas}$$

y

$$\sigma_{\bar{X}_A-\bar{X}_B} = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = \sqrt{\frac{(100)^2}{125} + \frac{(200)^2}{125}} = 20 \text{ horas}$$

La variable estandarizada correspondiente a la diferencia de medias es

$$Z = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_{\bar{X}_A-\bar{X}_B})}{\sigma_{\bar{X}_A-\bar{X}_B}} = \frac{(\bar{X}_A - \bar{X}_B) - 200}{20}$$

que tiene una distribución que se aproxima a una distribución normal.

a) La diferencia de 160 horas en unidades estándar $= (160 - 200)/20 = -2$.

$$\begin{aligned}\text{Probabilidad que se busca} &= (\text{área bajo la curva normal a la derecha de } z = -2)\\ &= 0.5000 + 0.4772 = 0.9772\end{aligned}$$

b) La diferencia de 250 horas en unidades estándar $= (250 - 200)/20 = 2.50$.

$$\begin{aligned}\text{Probabilidad que buscamos} &= (\text{área bajo la curva normal a la derecha de } z = 2.50)\\ &= 0.5000 - 0.4938 = 0.0062.\end{aligned}$$

5.14. Los balines de determinada marca pesan 0.50 onzas y su desviación estándar es de 0.02 onzas. ¿Cuál es la probabilidad de que la diferencia entre los pesos de dos lotes, cada uno de 1 000 balines, sea mayor a 2 onzas?

Sean $\bar{X}_1$ y $\bar{X}_2$ los pesos medios de los dos lotes. Entonces,

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = 0.50 - 0.50 = 0$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(0.02)^2}{1000} + \frac{(0.02)^2}{1000}} = 0.000895$$

La variable estandarizada correspondiente a la diferencia de medias es $Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - 0}{0.000895}$ que tiene una distribución muy cercana a la distribución normal.

La diferencia de 2 onzas entre los dos lotes es equivalente a la diferencia de $2/1\,000 = 0.002$ onzas en las medias. Esto se puede representar si $\bar{X}_1 - \bar{X}_2 \geq 0.002$ o $\bar{X}_1 - \bar{X}_2 \leq -0.002$, es decir,

$$Z \geq \frac{0.002 - 0}{0.000895} = 2.23 \qquad \text{o bien} \qquad Z \leq \frac{-0.002 - 0}{0.000895} = -2.23$$

Entonces,

$$P(Z \geq 2.23 \text{ o } Z \leq -2.23) = P(Z \geq 2.23) + P(Z \leq 2.23) = 2(0.5000 - 0.4871) = 0.0258$$

5.15. *A* y *B* juegan a cara o cruz, y cada uno lanza 50 monedas. *A* gana el juego si obtiene 5 caras más que *B*, si no, gana este último. Determinar las posibilidades en contra de que *A* gane un determinado juego.

Sean P_A y P_B las proporciones de caras obtenidas por *A* y *B*, respectivamente. Si suponemos que las monedas son legales, la probabilidad p de obtener una cara es $\frac{1}{2}$. Entonces,

$$\mu_{P_A-P_B} = \mu_{P_A} - \mu_{P_B} = 0$$

y

$$\sigma_{P_A-P_B} = \sqrt{\sigma_{P_A}^2 + \sigma_{P_B}^2} = \sqrt{\frac{pq}{n_A} + \frac{pq}{n_B}} = \sqrt{\frac{2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}{50}} = 0.10$$

La variable estandarizada correspondiente a la diferencia entre las proporciones es $Z = (P_A - P_B - 0)/0.10$.

Para considerar esta variable como continua, 5 o más caras significará 4.5 o más caras, así la diferencia entre las proporciones es $4.5/50 = 0.09$ o más, es decir, Z mayor o igual a $(0.09 - 0)/0.10 = 0.9$ (o $Z \geq 0.9$). La probabi-

lidad de que esto ocurra corresponde al área bajo la curva normal a la derecha de $Z = 0.9$, que es $0.5000 - 0.3159 = 0.1841$.

Por tanto, las probabilidades en contra de que A sea ganador son $(1 - 0.1841) : 0.1841 = 0.8159 : 0.1841$ o 4.43 a 1.

5.16. Las medidas de dos distancias son 27.3 pulgadas y 15.6 pulgadas, y sus desviaciones estándar (errores estándar) son 0.16 pulgadas y 0.08 pulgadas, respectivamente. Determinar la media y la desviación estándar de *a*) la suma, y *b*) la diferencia de las distancias.

Si se denotan estas distancias con D_1 y D_2, entonces

a)
$$\mu_{D_1+D_2} = \mu_{D_1} + \mu_{D_2} = 27.3 + 15.6 = 42.9 \text{ pulgadas}$$
$$\sigma_{D_1+D_2} = \sqrt{\sigma_{D_1}^2 + \sigma_{D_2}^2} = \sqrt{(0.16)^2 + (0.08)^2} = 0.18 \text{ pulgadas}$$

b)
$$\mu_{D_1-D_2} = \mu_{D_1} - \mu_{D_2} = 27.3 - 15.6 = 11.7 \text{ pulgadas}$$
$$\sigma_{D_1-D_2} = \sqrt{\sigma_{D_1}^2 + \sigma_{D_2}^2} = \sqrt{(0.16)^2 + (0.08)^2} = 0.18 \text{ pulgadas}$$

5.17. Unas bombillas eléctricas tienen una vida media de 1 500 horas y desviación estándar de 150 horas. Estas bombillas se conectan de manera que cuando una se funda, otra empiece a funcionar. Si las vidas medias están normalmente distribuidas, ¿cuál es la probabilidad de que la iluminación dure *a*) por lo menos 5 000 horas, *b*) cuando mucho 4 200 horas?

Si las vidas medias se denotan como L_1, L_2 y L_3. Entonces,

$$\mu_{L_1+L_2+L_3} = \mu_{L_1} + \mu_{L_2} + \mu_{L_3} = 1\,500 + 1\,500 + 1\,500 = 4\,500 \text{ horas}$$
$$\sigma_{L_1+L_2+L_3} = \sqrt{\sigma_{L_1}^2 + \sigma_{L_2}^2 + \sigma_{L_3}^2} = \sqrt{3(150)^2} = 260 \text{ horas}$$

a) 5 000 horas en unidades estándar $= (5\,000 - 4\,500) / 260 = 1.92$
Probabilidad que se busca = (área bajo la curva normal a la derecha de $z = 1.92$)
$= 0.5000 - 0.4726 = 0.0274$

b) 4 200 horas en unidades estándar $= (4\,200 - 4\,500) / 260 = -1.15$.
Probabilidad que se busca = (área bajo la curva normal a la derecha de $z = -1.15$)
$= 0.5\,000 - 0.3749 = 0.1251$

DISTRIBUCIÓN MUESTRAL DE VARIANZAS

5.18. Con los datos del problema 5.1, determinar *a*) la media de la distribución muestral de varianzas, *b*) la desviación estándar de la distribución muestral de varianzas, es decir, el error estándar de varianzas.

a) Las varianzas muestrales correspondientes a cada una de las 25 muestras del problema 5.1 son

0	0.25	4.00	9.00	20.25
0.25	0	2.25	6.25	16.00
4.00	2.25	0	1.00	6.25
9.00	6.25	1.00	0	2.25
20.25	16.00	6.25	2.25	0

La media de la distribución muestral de varianzas es

$$\mu_{S^2} = \frac{\text{suma de todas las varianzas de la tabla anterior}}{25} = \frac{135}{25} = 5.40$$

Para ilustrar el hecho de que $\mu_{S^2} = (n - 1)(\sigma^2)/n$, dado que $n = 2$ y $\sigma^2 = 10.8$ [vea el problema 5.1 *b*)], el lado derecho es $\frac{1}{2}(10.8) = 5.4$.

Este resultado indica que una varianza correcta de las muestras suele definirse como $\hat{S}^2 = \frac{n}{n-1}S^2$, ya que entonces se deduce que $\mu_{\hat{S}^2} = \sigma^2$ (vea también observaciones de la página 158).

b) La varianza de la distribución muestral de varianzas $\sigma^2_{S^2}$, la obtenemos restando la media 5.40 de cada uno de los 25 números de la tabla anterior, elevando al cuadrado el resultado, sumando estos cuadrados y después dividimos el resultado entre 25. Por tanto, $\sigma^2_{S^2} = 575.75/25 = 23.03$ o $\sigma_{S^2} = 4.80$.

5.19. Repetir el problema anterior para el caso en el que el muestreo se haga sin reemplazo.

a) Hay 10 muestras cuyas varianzas corresponden a los números sobre (o bajo) la diagonal de ceros en la tabla del problema 5.18*a*). Entonces,

$$\mu_{S^2} = \frac{0.25 + 4.00 + 9.00 + 20.25 + 2.25 + 6.25 + 16.00 + 1.00 + 6.25 + 2.25}{10} = 6.75$$

Éste es un caso especial del resultado general $\mu_{S^2} = \left(\frac{N}{N-1}\right)\left(\frac{n-1}{n}\right)\sigma^2$ [ecuación (19), página 158] como comprobamos al sustituir $N = 5$, $n = 2$ y $\sigma^2 = 10.8$, a la derecha, con lo que obtenemos $\mu_{S^2} = \left(\frac{5}{4}\right)\left(\frac{1}{2}\right)(10.8) = 6.75$.

b) Restando 6.75 de cada uno de los 10 números sobre la diagonal de ceros de la tabla del problema 5.18 *a*), se elevan al cuadrado estos números, se suman y se divide el resultado entre 10, se obtiene $\sigma^2_{S^2} = 39.675$, o $\sigma_{S^2} = 6.30$.

5.20. Demostrar que

$$E(S^2) = \frac{n-1}{n}\sigma^2$$

donde S^2 es la varianza muestral de una muestra aleatoria de tamaño n, según se definió en las páginas 157-158 y σ^2 es la varianza poblacional.

Método 1
Se tiene

$$X_1 - \bar{X} = X_1 - \frac{1}{n}(X_1 + \cdots + X_n) = \frac{1}{n}[(n-1)X_1 - X_2 - \cdots - X_n]$$

$$= \frac{1}{n}[(n-1)(X_1 - \mu) - (X_2 - \mu) - \cdots - (X_n - \mu)]$$

Entonces,

$$(X_1 - \bar{X})^2 = \frac{1}{n^2}[(n-1)^2(X_1 - \mu)^2 + (X_2 - \mu)^2 + \cdots + (X_n - \mu)^2 + \text{términos de productos cruzados}]$$

Como las X son independientes, la esperanza de cada término del producto cruzado es cero, con lo que tenemos

$$E[(X_1 - \bar{X})^2] = \frac{1}{n^2}\{(n-1)^2E[(X_1 - \mu)^2] + E[(X_2 - \mu)^2] + \cdots + E[(X_n - \mu)^2]\}$$

$$= \frac{1}{n^2}\{(n-1)^2\sigma^2 + \sigma^2 + \cdots + \sigma^2\}$$

$$= \frac{1}{n^2}\{(n-1)^2\sigma^2 + (n-1)\sigma^2\} = \frac{n-1}{n}\sigma^2$$

De manera similar, $E[(X_k - \bar{X})^2] = (n-1)\sigma^2/n$ para $k = 2, \ldots, n$. Por tanto,

$$E(S^2) = \frac{1}{n}E[(X_1 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2]$$

$$= \frac{1}{n}\left[\frac{n-1}{n}\sigma^2 + \cdots + \frac{n-1}{n}\sigma^2\right] = \frac{n-1}{n}\sigma^2$$

Método 2

Tenemos $X_j - \bar{X} = (X_j - \mu) - (\bar{X} - \mu)$. Entonces,

$$(X_j - \bar{X})^2 = (X_j - \mu)^2 - 2(X_j - \mu)(\bar{X} - \mu) + (\bar{X} - \mu)^2$$

y

$$\sum(X_j - \bar{X})^2 = \sum(X_j - \mu)^2 - 2(\bar{X} - \mu)\sum(X_j - \mu) + \sum(\bar{X} - \mu)^2 \tag{1}$$

donde la suma va desde $j = 1$ hasta n. Esto puede expresarse como

$$\begin{aligned}\sum(X_j - \bar{X})^2 &= \sum(X_j - \mu)^2 - 2n(\bar{X} - \mu)^2 + n(\bar{X} - \mu)^2 \\ &= \sum(X_j - \mu)^2 - n(\bar{X} - \mu)^2\end{aligned} \tag{2}$$

dado que $\sum(X_j - \mu) = \sum X_j - (n \cdot \mu) = n(\bar{X} - \mu)$. Si tomamos esperanzas a ambos lados de (2) y del problema 5.7, encontramos

$$E\left[\sum(X_j - \bar{X})^2\right] = E\left[\sum(X_j - \mu)^2\right] - nE[(\bar{X} - \mu)^2]$$

$$= n\sigma^2 - n\left(\frac{\sigma^2}{n}\right) = (n - 1)\sigma^2$$

de donde

$$E(S^2) = \frac{n - 1}{n}\sigma^2$$

5.21. Demostrar el teorema 5-4 de la página 156.

Si X_j, $j = 1, 2, \ldots, n$, está distribuida de manera normal con media μ y varianza σ^2, entonces su función característica es (vea la tabla 4-2, página 110)

$$\phi_j(\omega) = e^{i\mu\omega - (\sigma^2\omega^2)/2}$$

La función característica de $X_1 + X_2 + \cdots X_n$ es, entonces, de acuerdo con el teorema 3-12,

$$\phi(\omega) = \phi_1(\omega)\phi_2(\omega)\cdots\phi_n(\omega) = e^{in\mu\omega - (n\sigma^2\omega^2)/2}$$

dado que las X_j son independientes. De acuerdo con el teorema 3-11, la función característica de

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

es

$$\phi_{\bar{X}}(\omega) = \phi\left(\frac{\omega}{n}\right) = e^{i\mu\omega - [(\sigma^2/n)\omega^2]/2}$$

Pero ésta es la función característica de una distribución normal con media μ y varianza σ^2/n y el resultado que buscamos se deduce del teorema 3-13.

5.22. Demostrar el teorema 5-6 de la página 158.

Por definición $(n - 1)\hat{S}^2 = \sum_{j=1}^{n}(X_j - \bar{X})^2$. Entonces se deduce, de acuerdo con (2) del método 2 del problema 5.20, que $V = V_1 + V_2$, donde

$$V = \sum_{j=1}^{n}\frac{(X_j - \mu)^2}{\sigma^2}, \qquad V_1 = \frac{(n - 1)\hat{S}^2}{\sigma^2}, \qquad V_2 = \frac{(\bar{X} - \mu)^2}{\sigma^2/n}$$

Ahora, de acuerdo con el teorema 4-3, página 115, V tiene una distribución ji cuadrada con n grados de libertad [como observamos sustituyendo X_j por $(X_j - \mu)/\sigma$]. De acuerdo con el problema 5.21, $\bar{X}$ está distribuida normalmente con media μ y varianza σ^2/n. Por tanto, de acuerdo con el teorema 4-3 con $\nu = 1$ y X_1 sustituida por $(\bar{X} - \mu)/\sqrt{\sigma^2/n}$, vemos que V_2 es ji cuadrada con 1 grado de libertad. Del teorema 4-5, página 115, tenemos que si V_1 y V_2 son independientes, entonces V_1 tiene una distribución ji cuadrada con n – 1 grados de libertad. Dado que puede demostrarse que V_1 y V_2 son independientes, concluimos que es el resultado deseado.

5.23. *a*) Use el problema 5.6 para determinar, en el problema 5.1, la cantidad esperada de muestras en las que la varianza muestral será mayor a 7.2. *b*) Comprobar el resultado del inciso *a*) con el resultado verdadero.

a) Se tiene $n = 2$, $\sigma^2 = 10.8$ [de acuerdo con el problema 5.1*b*)]. Si $s_1^2 = 7.2$, tenemos

$$\frac{ns_1^2}{\sigma^2} = \frac{(2)(7.2)}{10.8} = 1.33$$

De acuerdo con el teorema 5-6, $\chi^2 = nS^2/\sigma^2 = 2S^2/10.8$ tiene distribución ji cuadrada con un grado de libertad. Consultando la tabla del apéndice E concluimos que

$$P(S^2 \geq s_1^2) = P(\chi^2 \geq 1.33) = 0.25$$

Por tanto, se espera que alrededor de 25% de las muestras, es decir 6, tengan una varianza mayor que 7.2.

b) De acuerdo con el problema 5.18, mediante conteo encontramos que en realidad hay 6 varianzas mayores a 7.2, de manera que hay coincidencia.

CASO EN EL QUE NO SE CONOCE LA VARIANZA POBLACIONAL

5.24. Demostrar el teorema 5-7 de la página 159.

Sea $Y = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$, $Z = \dfrac{nS^2}{\sigma^2}$, $\nu = n - 1$. Entonces, como las X_j están distribuidas de manera normal con media μ y varianza σ^2, se sabe (problema 5.21) que $\bar{X}$ está normalmente distribuida con media μ y varianza σ^2/n, de manera que y está distribuida normalmente con media 0 y varianza 1. De acuerdo con el teorema 5-6 de la página 158, o con el problema 5.22, Z tiene una distribución ji cuadrada con $\nu = n - 1$ grados de libertad. Además, se puede mostrar que Y y Z son independientes.

Del teorema 4-6 de la página 116 se infiere que

$$T = \frac{Y}{\sqrt{Z/\nu}} = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} = \frac{\bar{X} - \mu}{\hat{S}/\sqrt{n}}$$

tiene una distribución t con $n - 1$ grados de libertad.

5.25. De acuerdo con la tabla de la distribución t de Student para un grado de libertad (apéndice D), se tiene $P(-1.376 \leq T \leq 1.376) = 0.60$. Comprobar si este resultado se confirma con los resultados obtenidos en el problema 5.1.

De acuerdo con los valores de $\bar{X}$ en (1), en la página 155, y con los valores de S^2 en el problema 5.18*a*), obtenemos los valores siguientes para $T = (\bar{X} - \mu)/(S/\sqrt{1})$:

$-\infty$	-7.0	-1.0	-0.33	0.11
-7.0	$-\infty$	-1.0	-0.20	0.25
-1.0	-1.0	$\cdots$	1.0	1.0
-0.33	-0.20	1.0	∞	2.33
0.11	0.25	1.0	2.33	∞

En realidad existen 16 valores para los que $-1.376 \leq T \leq 1.376$ mientras que se deberían esperar $(0.60)(25) = 15$. Este resultado no es tan malo si se considera la pequeña cantidad de valores empleados. Este método de muestreo fue el que empleó "Student" para obtener originalmente la distribución t.

DISTRIBUCIÓN MUESTRAL DE RAZONES DE VARIANZAS

5.26. Demuestre el teorema 5-8 de la página 159.

Las muestras de tamaño m y n se denotan por $X_1, \ldots, X_n$ y $Y_1, \ldots, Y_n$, respectivamente. Entonces, las varianzas muestrales están dadas por

$$S_1^2 = \frac{1}{m}\sum_{j=1}^{m}(X_j - \bar{X})^2, \qquad S_2^2 = \frac{1}{n}\sum_{j=1}^{n}(Y_j - \bar{Y})^2$$

donde $\bar{X}$ y $\bar{Y}$ son las medias muestrales.

Ahora, de acuerdo con el teorema 5-6 de la página 158, se sabe que mS_1^2/σ_1^2 y nS_2^2/σ_2^2 tienen distribución ji cuadrada con $m - 1$ y $n - 1$ grados de libertad, respectivamente. Por tanto, de acuerdo con el teorema 4-7 de la página 117, concluimos que

$$F = \frac{mS_1^2/(m-1)\sigma_1^2}{nS_2^2/(n-1)\sigma_2^2} = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2}$$

tiene distribución F con $m - 1$ y $n - 1$ grados de libertad.

5.27. De dos poblaciones distribuidas normalmente cuyas varianzas son 20 y 36, respectivamente, se toman dos muestras de tamaño 8 y 10, respectivamente. Encontrar la probabilidad de que la varianza de la primera muestra sea más del doble de la varianza de la segunda.

Se tiene $m = 8, n = 10, \sigma_1^2 = 20, \sigma_2^2 = 36$. Por tanto,

$$F = \frac{8S_1^2/(7)(20)}{10S_2^2/(9)(36)} = 1.85\,\frac{S_1^2}{S_2^2}$$

El número de grados de libertad en el numerador y en el denominador es $\nu_1 = m - 1 = 7, \nu_2 = n - 1 = 9$. Ahora, si S_1^2 es más del doble de S_2^2, es decir, $S_1^2 > 2S_2^2$, entonces $F > 3.70$. Consultando las tablas del apéndice F, se ve que esta probabilidad es menor que 0.05, pero mayor que 0.01. Para obtener valores más exactos es necesario contar con una tabulación más amplia de la distribución F.

DISTRIBUCIONES DE FRECUENCIAS

5.28. En la tabla 5-4 se muestran los pesos, en libras, de 40 estudiantes varones de una universidad. Elaborar una distribución de frecuencias.

Tabla 5-4

138	164	150	132	144	125	149	157
146	158	140	147	136	148	152	144
168	126	138	176	163	119	154	165
146	173	142	147	135	153	140	135
161	145	135	142	150	156	145	128

El peso mayor es de 176 libras y el menor es de 119 libras, por lo que el rango es 176 – 119 = 57 lb.

Si se hacen 5 clases, el tamaño del intervalo de clase es $57/5 = 11$ aproximadamente.

Si se hacen 20 clases, el tamaño del intervalo de clase es $57/20 = 3$ aproximadamente.

Un tamaño adecuado para el intervalo de clase es 5 lb. Las marcas de clase serán 120, 125, 130, 135, . . . libras. Por tanto, los intervalos de clase serán 118-122, 123-127, 128-132, . . . Y los límites de clase serán 117.5, 122.5, 127.5, . . . , los cuales no coinciden con los datos observados.

La distribución de frecuencias que buscamos se muestra en la tabla 5-5. La columna central, llamada *columna de conteo*, que se usa para ingresar las frecuencias de clase que se encontraron de los datos en bruto, por lo general se omite en la presentación final de la distribución de frecuencia.

Otra posibilidad

Por supuesto, hay otras distribuciones de frecuencia. En la tabla 5-6, por ejemplo, se muestra una distribución de frecuencias en la que sólo hay 7 clases y en la que el intervalo de clase es de 9 libras.

Tabla 5-5

Peso (lb)	Conteo	Frecuencia
118-122	/	1
123-127	//	2
128-132	//	2
133-137	////	4
138-142	𝍸 /	6
143-147	𝍸 ///	8
148-152	𝍸	5
153-157	////	4
158-162	//	2
163-167	///	3
168-172	/	1
173-177	//	2
	TOTAL	40

Tabla 5-6

Peso (lb)	Conteo	Frecuencia
118-126	///	3
127-135	𝍸	5
136-144	𝍸 ////	9
145-153	𝍸 𝍸 //	12
154-162	𝍸	5
163-171	////	4
172-180	//	2
	TOTAL	40

5.29. Trazar un histograma y un polígono de frecuencia con la distribución de pesos del problema 5.28.

En las figuras 5-7 y 5-8 se presentan el histograma y el polígono de frecuencias correspondientes a cada uno de los dos casos considerados en el problema 5.28. Vemos que los centros de las bases de los rectángulos coinciden con las marcas de clase.

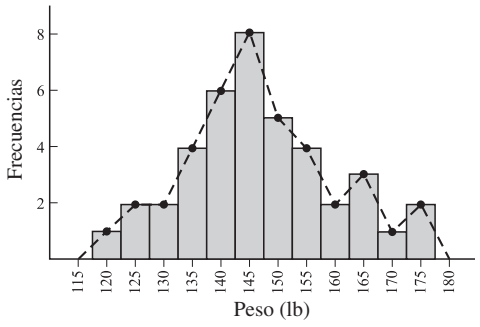


Figura 5-7

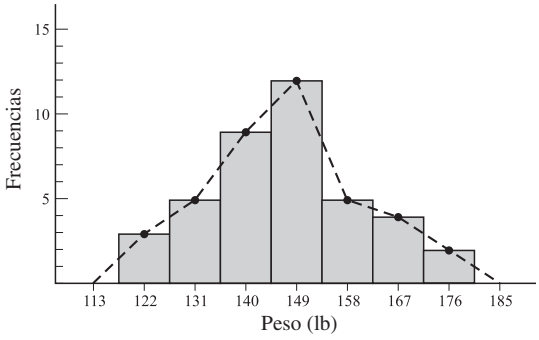


Figura 5-8

5.30. Cinco monedas se lanzan simultáneamente 1 000 veces y en cada lanzamiento se anota el número de caras obtenidas. En la tabla 5-7 se muestra el número de lanzamientos en los que se obtuvieron 0, 1, 2, 3, 4 y 5 caras. Representar gráficamente los datos.

Estos datos pueden representarse gráficamente ya sea como en la figura 5-9 o como en la figura 5-10.

La figura 5-9 parece ser una manera más natural de representar estos datos, lo cual se debe a que el número de caras no puede ser 1.5 o 3.2. Esta gráfica es un tipo de gráfica de barras de amplitud cero y suele llamársele *gráfica de líneas*. Se usa especialmente cuando se tienen datos discretos.

En la figura 5-10 se muestra un histograma de los datos. Observe que el área total del histograma es la frecuencia total 1 000, como debe ser.

Tabla 5-7

Cantidad de caras	Cantidad de lanzamientos (frecuencias)
0	38
1	144
2	342
3	287
4	164
5	25
TOTAL	1 000

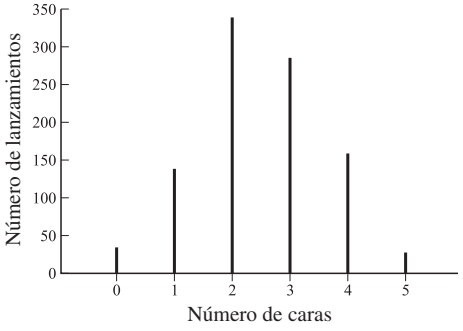


Figura 5-9

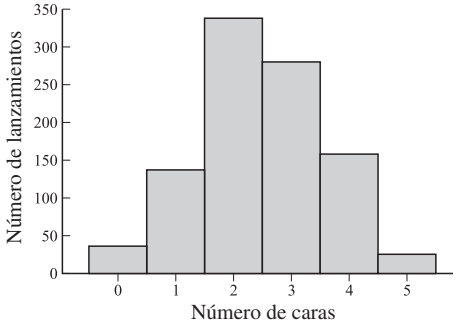


Figura 5-10

CÁLCULO DE LA MEDIA, VARIANZA Y MOMENTOS MUESTRALES

5.31. Encontrar la media aritmética de los números 5, 3, 6, 5, 4, 5, 2, 8, 6, 5, 4, 8, 3, 4, 5, 4, 8, 2, 5, 4.

Método 1

$$\bar{x} = \frac{\sum x}{n} = \frac{5 + 3 + 6 + 5 + 4 + 5 + 2 + 8 + 6 + 5 + 4 + 8 + 3 + 4 + 5 + 4 + 8 + 2 + 5 + 4}{20}$$

$$= \frac{96}{20} = 4.8$$

Método 2

Hay 6 cincos, 2 tres, 2 seis, 5 cuatros, 2 dos y 3 ochos. Entonces,

$$\bar{x} = \frac{\sum fx}{n} = \frac{(6)(5) + (2)(3) + (2)(6) + (5)(4) + (2)(2) + (3)(8)}{6 + 2 + 2 + 5 + 2 + 3} = \frac{96}{20} = 4.8$$

5.32. En cuatro grupos de estudiantes de 15, 20, 10 y 18 individuos el peso medio es 162, 148, 153 y 140 lb, respectivamente. Encontrar el peso medio de todos los estudiantes.

$$\bar{x} = \frac{\sum fx}{n} = \frac{(15)(162) + (20)(148) + (10)(153) + (18)(140)}{15 + 20 + 10 + 18} = 150 \text{ lb}$$

5.33. Con la distribución de frecuencias de las estaturas que se da en la tabla 5-2 de la página 161 hallar la estatura media de los 100 estudiantes varones de la universidad XYZ.

Observamos que en la tabla 5-8 se considera que todos los estudiantes con una estatura de 60-62 pulgadas, 63-65 pulgadas, etc., tienen las estaturas 61, 64, etc., pulgadas. Entonces, el problema se reduce a hallar la estatura media de 100 estudiantes si la estatura de 5 de ellos es de 61 pulgadas, la de 18 de ellos es de 64 pulgadas, etcétera.

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{\sum fx}{n} = \frac{6\,745}{100} = 67.45 \text{ pulgadas}$$

Tabla 5-8

Estatura (pulg)	Marca de clase (x)	Frecuencia (f)	fx
60-62	61	5	305
63-65	64	18	1 152
66-68	67	42	2 814
69-71	70	27	1 890
72-74	73	8	584
		$n = \sum f = 100$	$\sum fx = 6\,745$

Estos cálculos pueden ser tediosos, especialmente si se tiene que trabajar con números grandes y muchas clases. Para tales casos existen técnicas abreviadas (vea el problema 5.35).

5.34. Deducir la fórmula de compilación (27) de la página 162 para la media aritmética.

Sea x_j la marca de clase j-ésima. Entonces, la desviación de x_j de una determinada marca de clase a, es $x_j - a$, la cual será igual al tamaño del intervalo de clase c multiplicado por algún entero u_j, es decir, $x_j - a = cu_j$ o $x_j = a + cu_j$ (lo que también se escribe abreviado como $x = a + cu$).

Entonces, la media está dada por

$$\bar{x} = \frac{\sum f_j x_j}{n} = \frac{\sum f_j(a + cu_j)}{n} = \frac{a\sum f_j}{n} + c\frac{\sum f_j u_j}{n}$$
$$= a + c\frac{\sum f_j u_j}{n} = a + c\bar{u}$$

ya que $n = \sum f_j$.

5.35. Use la fórmula de compilación del problema 5.34 para hallar la estatura media de los 100 estudiantes varones de la universidad XYZ (vea el problema 5.33).

Para facilitar los cálculos se elabora una tabla como la 5-9. Este método se llama *método de compilación* y es recomendable usarlo siempre que sea posible.

$$\bar{x} = a + \left(\frac{\sum fu}{n}\right)c = 67 + \left(\frac{15}{100}\right)(3) = 67.45 \text{ pulgadas}$$

Tabla 5-9

	x	u	f	fu
	61	−2	5	−10
	64	−1	18	−18
$a \rightarrow$	67	0	42	0
	70	1	27	27
	73	2	8	16
			$n = 100$	$\sum fu = 15$

5.36. Calcular *a*) la varianza, *b*) la desviación estándar para los números del problema 5.31.

a) **Método 1**
Como en el problema 5.31, tenemos $\bar{x} = 4.8$. Entonces,

$$s^2 = \frac{\sum(x - \bar{x})^2}{n} = \frac{(5 - 4.8)^2 + (3 - 4.8)^2 + (6 - 4.8)^2 + (5 - 4.8)^2 + \cdots + (4 - 4.8)^2}{20}$$

$$= \frac{59.20}{20} = 2.96$$

Método 2

$$s^2 = \frac{\sum f(x - \bar{x})^2}{n} = \frac{6(5 - 4.8)^2 + 2(3 - 4.8)^2 + 2(6 - 4.8)^2 + 5(4 - 4.8)^2 + 3(8 - 4.8)^2}{20}$$

$$= \frac{59.20}{20} = 2.96$$

b) De acuerdo con el inciso *a*), $s^2 = 2.96$ y $s = \sqrt{2.96} = 1.72$.

5.37. Determinar la desviación estándar de los pesos de los estudiantes del problema 5.32.

$$s^2 = \frac{\sum f(x - \bar{x})^2}{n} = \frac{15(162 - 150)^2 + 20(148 - 150)^2 + 10(153 - 150)^2 + 18(140 - 150)^2}{15 + 20 + 10 + 18}$$

$$= \frac{4\,130}{63} = 65.6 \text{ en libras al cuadrado, o bien (lb)}^2$$

Entonces, $s = \sqrt{65.6\ (\text{lb})^2} = \sqrt{65.6}$ lb = 8.10 lb, donde nos hemos basado en el hecho de que las unidades siguen las leyes usuales del álgebra.

5.38. Calcular la desviación estándar de las estaturas de los 100 estudiantes de la universidad XYZ (vea el problema 5.33). De acuerdo con el problema 5.33, $\bar{X} = 67.45$ pulgadas. Para facilitar los cálculos elaboramos una tabla como la 5-10.

$$s = \sqrt{\frac{\sum f(x - \bar{x})^2}{n}} = \sqrt{\frac{852.7500}{100}} = \sqrt{8.5275} = 2.92 \text{ pulgadas}$$

Tabla 5-10

Estatura (pulgada)	Marca de clase (x)	$x - \bar{x} = x - 67.45$	$(x - \bar{x})^2$	Frecuencias (f)	$f(x - \bar{x})^2$
60-62	61	−6.45	41.6025	5	208.0125
63-65	64	−3.45	11.9025	18	214.2450
66-68	67	−0.45	0.2025	42	8.5050
69-71	70	2.55	6.5025	27	175.5675
72-74	73	5.55	30.8025	8	246.4200
				$n = \sum f = 100$	$\sum f(x - \bar{x})^2 = 852.7500$

5.39. Deducir la fórmula de compilación (28), de la página 162, para la varianza.

Como en el problema 5.34, tenemos $x_j = a + cu_j$ y

$$\bar{x} = a + c\frac{\sum f_j u_j}{n} = a + c\bar{u}$$

Entonces,

$$s^2 = \frac{1}{n}\sum f_j(x_j - \bar{x})^2 = \frac{1}{n}\sum f_j(cu_j - c\bar{u})^2$$

$$= \frac{c^2}{n}\sum f_j(u_j - \bar{u})^2$$

$$= \frac{c^2}{n}\sum f_j(u_j^2 - 2u_j\bar{u} + \bar{u}^2)$$

$$= \frac{c^2}{n}\sum f_j u_j^2 - \frac{2\bar{u}c^2}{n}\sum f_j u_j + \frac{c^2}{n}\sum f_j\bar{u}^2$$

$$= c^2\frac{\sum f_j u_j^2}{n} - 2\bar{u}^2c^2 + c^2\bar{u}^2$$

$$= c^2\frac{\sum f_j u_j^2}{n} - c^2\left(\frac{\sum f_j u_j}{n}\right)^2$$

$$= c^2\left[\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2\right]$$

$$= c^2[\bar{u}^2 - \bar{u}^2]$$

5.40. Con la fórmula de compilación del problema 5.39 calcular la desviación estándar de las estaturas del problema 5.33.

Para facilitar los cálculos formamos una tabla como la 5-11. Esto permite calcular $\bar{x}$ como en el problema 5.35. Empleando la última columna tenemos

$$s^2 = c^2\left[\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2\right] = c^2(\bar{u}^2 - \bar{u}^2)$$

$$= (3)^2\left[\frac{97}{100} - \left(\frac{15}{100}\right)^2\right] = 8.5275$$

y, por tanto, $s = 2.92$ pulgadas.

Tabla 5-11

	x	u	f	fu	fu^2
	61	−2	5	−10	20
	64	−1	18	−18	18
$a \rightarrow$	67	0	42	0	0
	70	1	27	27	27
	73	2	8	8	32
			$n = \sum f = 100$	$\sum fu = 15$	$\sum fu^2 = 97$

5.41. Dada la distribución de estaturas del problema 5.33, encontrar los cuatro primeros momentos respecto a la media.

Continuando con el método del problema 5.40, formamos la tabla 5-12. Después, mediante el empleo de la notación de la página 162 tenemos

$$M_1' = \frac{\sum fu}{n} = 0.15 \qquad M_3' = \frac{\sum fu^3}{n} = 0.33$$

$$M_2' = \frac{\sum fu^2}{n} = 0.97 \qquad M_4' = \frac{\sum fu^4}{n} = 2.53$$

Tabla 5-12

x	u	f	fu	fu^2	fu^3	fu^4
61	−2	5	−10	20	−40	80
64	−1	18	−18	18	−18	18
67	0	42	0	0	0	0
70	1	27	27	27	27	27
73	2	8	16	32	64	128
		$n = \sum f = 100$	$\sum fu = 15$	$\sum fu^2 = 97$	$\sum fu^3 = 33$	$\sum fu^4 = 253$

y de (32),

$$m_1 = 0$$
$$m_2 = c^2\left(M_2' - M_1'^2\right) = 9[0.97 - (0.15)^2] = 8.5275$$
$$m_3 = c^3\left(M_3' - 3M_1'M_2' + 2M_1'^3\right) = 27[0.33 - 3(0.15)(0.97) + 2(0.15)^3] = -2.6932$$
$$m_4 = c^4\left(M_4' - 4M_1'M_3' + 6M_1'^2M_2' - 3M_1'^4\right)$$
$$= 81[2.53 - 4(0.15)(0.33) + 6(0.15)^2(0.97) - 3(0.15)^4] = 199.3759$$

5.42. Calcular los coeficientes de *a*) sesgo, *b*) curtosis de la distribución de estaturas del problema 5.33.

a) De acuerdo con el problema 5.41,

$$m_2 = s^2 = 8.5275 \qquad m_3 = -2.6932$$

Entonces,
$$\text{Coeficiente de sesgo} = a_3 = \frac{m_3}{s^3}$$
$$= \frac{-2.6932}{\sqrt{(8.5275)^3}} = -0.14$$

b) De acuerdo con el problema 5.41,

$$m_4 = 199.3759 \qquad m_2 = s^2 = 8.5275$$

Entonces,
$$\text{Coeficiente de curtosis} = a_4 = \frac{m_4}{s^4}$$
$$= \frac{199.3759}{(8.5275)^2} = 2.74$$

De acuerdo con el inciso *a*) vemos que la distribución es moderadamente sesgada a la izquierda. Del inciso *b*) notamos que esta distribución es ligeramente menos afilada que la distribución normal (cuyo coeficiente de curtosis es = 3).

PROBLEMAS DIVERSOS

5.43. *a*) Mostrar cómo seleccionar 30 muestras aleatorias, cada una de 4 estudiantes (con reemplazo), de la tabla de las estaturas de la página 161, usando números aleatorios. *b*) Encuentre la media y desviación estándar de la distribución muestral de medias del inciso *a*). *c*) Compare los resultados del inciso *b*) con los valores teóricos, y explicar las diferencias.

a) Para enumerar a los 100 estudiantes se emplean dos dígitos: 00, 01, 02, . . . , 99 (vea la tabla 5-13). Por tanto, los 5 estudiantes cuyas estaturas pertenecen a la categoría de 60-62 pulgadas se enumeran 00-04, los 18 cuyas estaturas pertenecen a la de 63-65 pulgadas se enumeran 05-22, etc. Al número de cada estudiante se le llama *número muestral*, o *de muestreo*.

Tabla 5-13

Estatura (pulgadas)	Frecuencia	Número de muestreo
60-62	5	00-04
63-65	18	05-22
66-68	42	23-64
69-71	27	65-91
72-74	8	92-99

Ahora se extraen números de muestreo de la tabla de números aleatorios (apéndice H). En el primer renglón se encuentra la secuencia 51, 77, 27, 46, 40, etc., que se consideran como números muestrales aleatorios, cada uno de los cuales da la estatura de un estudiante determinado. Por tanto, 51 corresponde a un estudiante que tiene la estatura 66-68 pulgadas, que se toma como 67 pulgadas (la marca de clase). De manera similar 77, 27, 46 dan las estaturas 70, 67, 67 pulgadas, respectivamente.

Mediante este proceso se obtiene la tabla 5-14, que muestra los números muestrales extraídos, las estaturas correspondientes y la estatura media de cada una de las 30 muestras. Se debe destacar que, aunque hemos utilizado los números aleatorios del primer renglón de la tabla de números aleatorios, podemos empezar con *cualquiera* y seguir cualquier patrón.

Tabla 5-14

Extracción de números muestrales		Estaturas correspondientes	Estatura media	Extracción de números muestrales		Estaturas correspondientes	Estatura media
1.	51, 77, 27, 46	67, 70, 67, 67	67.75	**16.**	11, 64, 55, 58	64, 67, 67, 67	66.25
2.	40, 42, 33, 12	67, 67, 67, 64	66.25	**17.**	70, 56, 97, 43	70, 67, 73, 67	69.25
3.	90, 44, 46, 62	70, 67, 67, 67	67.75	**18.**	74, 28, 93, 50	70, 67, 73, 67	69.25
4.	16, 28, 98, 93	64, 67, 73, 73	69.25	**19.**	79, 42, 71, 30	70, 67, 70, 67	68.50
5.	58, 20, 41, 86	67, 64, 67, 70	67.00	**20.**	58, 60, 21, 33	67, 67, 64, 67	66.25
6.	19, 64, 08, 70	64, 67, 64, 70	66.25	**21.**	75, 79, 74, 54	70, 70, 70, 67	69.25
7.	56, 24, 03, 32	67, 67, 61, 67	65.50	**22.**	06, 31, 04, 18	64, 67, 61, 64	64.00
8.	34, 91, 83, 58	67, 70, 70, 67	68.50	**23.**	67, 07, 12, 97	70, 64, 64, 73	67.75
9.	70, 65, 68, 21	70, 70, 70, 64	68.50	**24.**	31, 71, 69, 88	67, 70, 70, 70	69.25
10.	96, 02, 13, 87	73, 61, 64, 70	67.00	**25.**	11, 64, 21, 87	64, 67, 64, 70	66.25
11.	76, 10, 51, 08	70, 64, 67, 64	66.25	**26.**	03, 58, 57, 93	61, 67, 67, 73	67.00
12.	63, 97, 45, 39	67, 73, 67, 67	68.50	**27.**	53, 81, 93, 88	67, 70, 73, 70	70.00
13.	05, 81, 45, 93	64, 70, 67, 73	68.50	**28.**	23, 22, 96, 79	67, 64, 73, 70	68.50
14.	96, 01, 73, 52	73, 61, 70, 67	67.75	**29.**	98, 56, 59, 36	73, 67, 67, 67	68.50
15.	07, 82, 54, 24	64, 70, 67, 67	67.00	**30.**	08, 15, 08, 84	64, 64, 64, 70	65.50

b) En la tabla 5-15 se proporciona la distribución de frecuencias de las estaturas medias muestrales obtenidas en el inciso *a*). Ésta es una *distribución muestral de medias*. La media y la desviación estándar se obtienen empleando el método de compilación antes descrito.

$$\text{Media} = a + c\bar{u} = a + \frac{c\sum fu}{n} = 67.00 + \frac{(0.75)(23)}{30} = 67.58 \text{ pulgadas}$$

$$\text{Desviación estándar} = c\sqrt{\bar{u}^2 - \bar{u}^2} = 0\ c\sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2}$$

$$= (0.75)\sqrt{\frac{123}{30} - \left(\frac{23}{30}\right)^2} = 1.41 \text{ pulgadas}$$

Tabla 5-15

Media muestral	Conteo	f	u	fu	fu^2
64.00	/	1	−4	−4	16
64.75		0	−3	0	0
65.50	//	2	−2	−4	8
66.25	𝍸 /	6	−1	−6	6
$a \rightarrow$ 67.00	////	4	0	0	0
67.75	////	4	1	4	4
68.50	𝍸 //	7	2	14	28
69.25	𝍸	5	3	15	45
70.00	/	1	4	4	16
		$\sum f = n = 30$		$\sum fu = 23$	$\sum fu^2 = 123$

c) La media teórica de la distribución muestral de medias dada por $\mu_{\bar{X}}$, es igual a la media poblacional μ, que es de 67.45 pulgadas (vea el problema 5.33), lo cual coincide de manera estrecha con el valor 67.58 pulgadas del inciso *b*).

La desviación estándar teórica (error estándar) de la distribución muestral de medias, dada por $\sigma_{\bar{X}}$, es igual a $\sigma/\sqrt{n}$, donde la desviación estándar poblacional es $\sigma = 2.92$ pulgadas (vea el problema 5.40) y el tamaño de la muestra es $n = 4$. Dado que $\sigma/\sqrt{n} = 2.92/\sqrt{4} = 1.46$ pulgadas, coincide estrechamente con el valor 1.41 pulgadas del inciso *b*). Las discrepancias se deben a que sólo se tomaron 30 muestras y a que el tamaño de la muestra es pequeño.

5.44. La desviación estándar de los pesos de una población muy grande de estudiantes es de 10.0 libras. De esta población se toman muestras cada una de 200 estudiantes y se calcula la desviación estándar de los pesos en cada muestra. Encontrar *a*) la media, *b*) la desviación estándar de la distribución muestral de desviaciones estándar.

Consideramos que el muestreo se hace de una población infinita o de una población finita con reemplazo. De acuerdo con la tabla 5-1 de la página 160, tenemos:

a) $$\mu_S = \sigma = 10.0 \text{ libras}$$

b) $$\sigma_S = \frac{\sigma}{\sqrt{2n}} = \frac{10}{\sqrt{400}} = 0.50 \text{ libras}$$

5.45. ¿Qué porcentaje de las muestras del problema 5.44 tendrán una desviación estándar *a*) mayor que 11 libras, *b*) menor que 8.8 libras?

La distribución muestral de las desviaciones estándar es aproximadamente normal, su media es 10.0 libras y su desviación estándar 0.50 libras.

a) 11.0 libras en unidades estándar = $(11.0 - 10.0)/0.50 = 2.0$. Área bajo la curva normal a la derecha de $z = 2.0$ es $(0.5 - 0.4772) = 0.0228$; por tanto, el porcentaje que se requiere es 2.3%.

b) 8.8 libras en unidades estándar = $(8.8 - 10.0)/0.50 = -2.4$. El área bajo la curva normal a la izquierda de $z = -2.4$ es $(0.5 - 0.4918) = 0.0082$; por tanto, el porcentaje que buscamos es 0.8%.

5.46. De una población continua se toma una muestra aleatoria de 6 observaciones. ¿Cuál es la probabilidad de que por lo menos 2 de ellas sean menores que las primeras 4?

Si suponemos que la función de densidad de población es $f(x)$. La probabilidad de que 3 de las primeras 4 observaciones sean mayores que u, mientras que la cuarta se encuentre entre u y $u + du$ está dada por

(1) $$ {}_4C_3\left[\int_u^\infty f(x)\,dx\right]^3 f(u)\,du$$

La probabilidad de que las 2 últimas observaciones sean menores que u (y, por tanto, menores que las primeras 4) es

$$\left[\int_{-\infty}^{u} f(x)\,dx\right]^2 \tag{2}$$

Entonces, la probabilidad de que las primeras 4 sean mayores que u y las 2 últimas sean menores que u es el producto de (1) y (2), es decir,

$${}_4C_3\left[\int_{u}^{\infty} f(x)\,dx\right]^3 f(u)\,du\left[\int_{-\infty}^{u} f(x)\,dx\right]^2 \tag{3}$$

Dado que u puede tomar valores entre $-\infty$ e ∞, la probabilidad total de que por lo menos 2 observaciones sean menores que las primeras 4 es la integral de (3) desde $-\infty$ hasta ∞, es decir,

$${}_4C_3\int_{-\infty}^{\infty}\left[\int_{u}^{\infty} f(x)\,dx\right]^3\left[\int_{-\infty}^{u} f(x)\,dx\right]^2 f(u)\,du \tag{4}$$

Para evaluar esto, sea

$$v = \int_{-\infty}^{u} f(x)\,dx \tag{5}$$

Entonces,

$$dv = f(u)\,du \qquad 1 - v = \int_{u}^{\infty} f(x)\,dx \tag{6}$$

Cuando $u = \infty$, $v = 1$ y cuando $u = -\infty$, $v = 0$. Por tanto, (4) se transforma en

$${}_4C_3\int_0^1 v^2(1 - v)^3 dv = 4\frac{\Gamma(3)\Gamma(4)}{\Gamma(7)} = \frac{1}{15}$$

que es la probabilidad que buscamos. Es interesante observar que la probabilidad no depende de la distribución de probabilidad $f(x)$. Éste es un ejemplo de *estadística no paramétrica* ya que no es necesario conocer el parámetro poblacional.

Otro método

Las observaciones se denotan $x_1, x_2, \ldots, x_6$. Como la población es continua, suponemos que las x_i son distintas. Existen 6! maneras de acomodar los subíndices 1, 2, . . . , 6, y cualquiera de estas maneras tiene la misma posibilidad de ordenar las x_i para que estén en orden creciente. De estos 6! acomodos, en exactamente $4! \times 2!$ acomodos, x_1, x_2, x_3, x_4, serán las 4 observaciones más pequeñas y x_5, x_6 serán las 2 observaciones más grandes. La probabilidad que buscamos es, entonces,

$$\frac{4! \times 2!}{6!} = \frac{1}{15}$$

5.47. Sea $\{X_1, X_2, \ldots, X_n\}$ una muestra aleatoria de tamaño n extraída sin reemplazo de una población finita de tamaño N. Demostrar que si la media y la varianza poblacionales son μ y σ^2, entonces *a*) $E(X_j) = \mu$, *b*) $\mathrm{Cov}(X_j, X_k) = -\sigma^2/(N-1)$.

Suponemos que la población consta de un conjunto de números $(\alpha_1, \alpha_2, \ldots, \alpha_N)$, en el que las α no necesariamente son distintas. Un procedimiento de muestreo aleatorio es aquel en el que cada muestra n, de las α extraídas de N, tiene la misma probabilidad (es decir, $1/{}_N C_n$). Esto significa que las X_j están distribuidas de manera idéntica:

$$X_j = \begin{cases} \alpha_1 & \text{prob. } 1/N \\ \alpha_2 & \text{prob. } 1/N \\ \vdots & \\ \alpha_N & \text{prob. } 1/N \end{cases} \qquad (j = 1, 2, \ldots, n)$$

Sin embargo, no son mutuamente independientes. En efecto, para $j \neq k$, la distribución conjunta de X_j y X_k es la dada por

$$P(X_j = \alpha_\lambda, X_k = \alpha_\nu) = P(X_j = \alpha_\lambda)P(X_k = \alpha_\nu | X_j = \alpha_\lambda)$$

$$= \frac{1}{N} P(X_k = \alpha_\nu | X_j = \alpha_\lambda)$$

$$= \begin{cases} \frac{1}{N}\left(\frac{1}{N-1}\right) & \lambda \neq \nu \\ 0 & \lambda = \nu \end{cases}$$

donde λ y ν van desde 1 hasta N.

a)
$$E(X_j) = \sum_{\lambda=1}^{N} \alpha_\lambda P(X_j = \alpha_\lambda) = \frac{1}{N}\sum_{\lambda=1}^{N} \alpha_\lambda = \mu$$

b)
$$\text{Cov}(X_j, X_k) = E[X_j - \mu)(X_k - \mu)]$$

$$= \sum_{\lambda=1}^{N}\sum_{\nu=1}^{N} (\alpha_\lambda - \mu)(\alpha_\nu - \mu)P(X_j = \alpha_\lambda, X_k = \alpha_\nu)$$

$$= \frac{1}{N}\left(\frac{1}{N-1}\right)\sum_{\lambda \neq \nu = 1}^{N} (\alpha_\lambda - \mu)(\alpha_\nu - \mu)$$

donde la última suma tiene $N(N - 1)$ términos en total, que corresponden a todos los pares posibles de λ y ν no iguales.

Ahora, de acuerdo con el álgebra elemental,

$$[(\alpha_1 - \mu) + (\alpha_2 - \mu) + \cdots + (\alpha_N - \mu)]^2 = \sum_{\lambda=1}^{N} (\alpha_\lambda - \mu)^2 + \sum_{\lambda \neq \nu = 1}^{N} (\alpha_\lambda - \mu)(\alpha_\nu - \mu)$$

En esta ecuación, cuyo lado izquierdo es cero, por definición

$$\alpha_1 + \alpha_2 + \cdots + \alpha_N = N\mu$$

y la primera suma del lado derecho es igual, por definición, a $N\sigma^2$. Por tanto,

$$\sum_{\lambda \neq \nu = 1}^{N} (\alpha_\lambda - \mu)(\alpha_\nu - \mu) = -N\sigma^2$$

y

$$\text{Cov}(X_j, X_k) = \frac{1}{N}\left(\frac{1}{N-1}\right)(-N\sigma^2) = -\frac{\sigma^2}{N-1}$$

5.48. Probar que en el problema 5.47 *a*) la media, *b*) la varianza de la media muestral están dadas, respectivamente, por

$$\mu_{\bar{X}} = \mu \qquad \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

a)
$$E(\bar{X}) = E\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n}[E(X_1) + \cdots + E(X_n)]$$

$$= \frac{1}{n}(\mu + \cdots + \mu) = \mu$$

donde hemos usado el problema 5.47*a*).

b) Con base en los teoremas 3-5 y 3-16 (generalizados) y el problema 5.47, obtenemos

$$\text{Var}(\bar{X}) = \frac{1}{n^2}\text{Var}\left(\sum_{j=1}^{n} X_j\right) = \frac{1}{n^2}\left[\sum_{j=1}^{n}\text{Var}(X_j) + \sum_{j\neq k=1}^{n}\text{Cov}(X_j, X_k)\right]$$

$$= \frac{1}{n^2}\left[n\sigma^2 + n(n-1)\left(-\frac{\sigma^2}{N-1}\right)\right]$$

$$= \frac{\sigma^2}{n}\left[1 - \frac{n-1}{N-1}\right] = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

PROBLEMAS SUPLEMENTARIOS

DISTRIBUCIÓN MUESTRAL DE MEDIAS

5.49. Una población consta de cuatro números 3, 7, 11, 15. Considere todas las muestras posibles de tamaño dos que se puedan extraer con reemplazo de esta población. Calcule *a*) la media poblacional, *b*) la desviación estándar poblacional, *c*) la media de la distribución muestral de medias, *d*) la desviación estándar de la distribución muestral de medias. Compruebe los incisos *c*) y *d*) directamente a partir de los incisos *a*) y *b*) con base en las fórmulas adecuadas.

5.50. Resuelva el problema 5.49 si el muestreo se hace sin reemplazo.

5.51. Los pesos de 1 500 balines están distribuidos normalmente y su media es 22.40 onzas y su desviación estándar 0.048 onzas. Si de esta población se toman 300 muestras aleatorias de tamaño 36, determine la media y la desviación estándar esperadas de la distribución muestral de medias si el muestreo se hace *a*) con reemplazo, *b*) sin reemplazo.

5.52. Resuelva el problema 5.51 si la población consta de 72 balines.

5.53. En el problema 5.51, ¿la media de cuántas muestras aleatorias *a*) estará entre 22.39 y 22.41 onzas, *b*) será mayor a 22.42 onzas, *c*) será menor a 22.37 onzas, *d*) menor que 22.38 onzas o mayor a 22.41 onzas?

5.54. Los cinescopios fabricados por una determinada empresa tienen una vida media de 800 horas con una desviación estándar de 60 horas. Determine la probabilidad de que una muestra aleatoria de 16 cinescopios tomados del grupo tenga una vida media *a*) entre 790 y 810 horas, *b*) menor a 785 horas, *c*) mayor a 820 horas, *d*) entre 770 y 830 horas.

5.55. Repita el problema 5.54 si la muestra es de 64 cinescopios. Explique la diferencia.

5.56. Los pesos de los paquetes que se reciben en una tienda departamental tienen una media de 300 libras y una desviación estándar de 50 libras. ¿Cuál es la probabilidad de que 25 paquetes recibidos al azar excedan el límite de seguridad del elevador que es de 8 200 libras?

DISTRIBUCIÓN MUESTRAL DE PROPORCIONES

5.57. Encuentre la probabilidad que de los siguientes 200 nacimientos *a*) menos de 40% sean niños, *b*) entre 43 y 57% sean niñas, *c*) más de 54% sean niños. Suponga la misma probabilidad para el nacimiento de un niño o de una niña.

5.58. En 1 000 muestras de 200 infantes cada una, ¿en cuántas puede esperarse encontrar que *a*) menos de 40% sean niños, *b*) entre 40 y 60% sean niñas, *c*) 53% o más sean niñas?

5.59. Repita el problema 5.57 si en lugar de 200 infantes se consideran 100 infantes, y explique las diferencias entre los resultados.

5.60. Una urna contiene 80 canicas de las cuales 60% son rojas y 40% blancas. En 50 muestras de 20 canicas, cada una tomada con reemplazo, devolución a la urna, ¿cuántas muestras puede esperarse que contengan *a*) el mismo número de canicas rojas y blancas, *b*) 12 canicas rojas y 8 blancas, *c*) 8 canicas rojas y 12 blancas, *d*) 10 o más canicas blancas?

5.61. Diseñe un experimento que tenga como objetivo ilustrar los resultados del problema 5.60. En lugar de canicas rojas y blancas pueden emplearse papelillos en los que estén escritas las letras *R* o *B*, en la proporción correcta. ¿Qué error puede introducirse al usar dos conjuntos diferentes de canicas?

5.62. Un fabricante tiene 1 000 lotes, cada uno de 100 bombillas eléctricas. Si normalmente 5% de éstas tienen algún defecto, ¿en cuántos lotes se esperará *a*) menos de 90 bombillas buenas, *b*) 98 o más bombillas buenas?

DISTRIBUCIONES MUESTRALES DE DIFERENCIAS Y SUMAS

5.63. La resistencia media a la ruptura de los cables producidos por los fabricantes *A* y *B* es de 4 000 y 4 500 libras con desviaciones estándar de 300 y 200 libras, respectivamente. Si se prueban 100 cables de *A* y 50 cables de *B*, ¿cuál es la probabilidad de que la resistencia media a la ruptura de los cables de *B* sea *a*) por lo menos 600 libras mayor que la de los cables de *A*, *b*) por lo menos 450 libras mayor que la resistencia de los de *A*?

5.64. En el problema 5.63, ¿cuál es la probabilidad en el caso de que se prueben 100 cables de cada marca? Explique las diferencias.

5.65. En una prueba de aptitudes, la calificación media de los estudiantes fue de 72 puntos y la desviación estándar de 8 puntos. ¿Cuál es la probabilidad de que dos grupos de estudiantes, uno de 28 y otro de 36, difieran en las medias de sus calificaciones en *a*) 3 o más puntos, *b*) 6 o más puntos, *c*) entre 2 y 5 puntos?

5.66. Una urna contiene 60 canicas rojas y 40 blancas. Se extraen con reemplazo dos conjuntos de 30 canicas cada una y se anota su color. ¿Cuál es la probabilidad de que la diferencia entre los dos conjuntos sea de 8 o más canicas rojas?

5.67. Resuelva el problema 5.66 si los muestreos se hacen sin reemplazo.

5.68. Como resultado de una elección, un candidato obtiene 65% de los votos. Determine la probabilidad de que en dos muestras aleatorias, cada una de 200 votantes, se encuentre una diferencia de 10% en la proporción de votantes a favor del candidato.

5.69. Si U_1 y U_2 son los conjuntos de números del problema 5.12, compruebe que *a*) $\mu_{U_1+U_2} = \mu_{U_1} + \mu_{U_2}$, *b*) $\sigma_{U_1+U_2} = \sqrt{\sigma_{U_1}^2 + \sigma_{U_2}^2}$.

5.70. Se tienen tres pesos de 20.48, 35.97 y 62.34 libras con desviaciones estándar de 0.21, 0.46 y 0.54 lb, respectivamente. Calcule *a*) la media y *b*) la desviación estándar de la suma de estos pesos.

5.71. El voltaje de una batería tiene una distribución muy cercana a la normal con media de 15.0 voltios y desviación estándar de 0.2 voltios. ¿Cuál es la probabilidad de que cuatro baterías conectadas en serie tengan un voltaje conjunto de 60.8 o más voltios?

DISTRIBUCIÓN MUESTRAL DE VARIANZAS

5.72. Con los datos del problema 5.49, determine *a*) la media de la distribución muestral de las varianzas, *b*) el error muestral de las varianzas.

5.73. Repita el problema 5.72 para el caso en el que el muestreo se hace sin reemplazo.

5.74. La varianza de una población normal es 15. Si de esta población se toman muestras de tamaño 5, ¿qué porcentaje puede esperarse que tenga una varianza *a*) menor que 10, *b*) mayor que 20, *c*) entre 5 y 10?

5.75. Se encuentra que las vidas medias de los cinescopios fabricados por una compañía tienen una distribución normal cuya media es de 2 000 horas y desviación estándar de la muestra es de 60 horas. Si se toman al azar 10 cinescopios, determine la probabilidad de que la desviación estándar *a*) no sea mayor a 50 horas, *b*) esté entre 50 y 70 horas.

CASO EN EL QUE NO SE CONOCE LA VARIANZA POBLACIONAL

5.76. De acuerdo con la tabla de la distribución t de Student para un grado de libertad (apéndice D), se tiene $P(-1 \leq T \leq 1) = 0.50$. Verifique si los resultados del problema 5.1 son confirmados por este valor, y explique cualquier diferencia.

5.77. Verifique si los resultados del problema 5.49 se confirman empleando *a*) $P(-1 \leq T \leq 1) = 0.50$, *b*) $P(-1.376 \leq T \leq 1.376) = 0.60$, donde T tiene la distribución t de Student con $\nu = 1$.

5.78. Explique cómo puede usarse el teorema 5-7 de la página 159 para elaborar una tabla de distribución t de Student como la del apéndice D.

DISTRIBUCIÓN MUESTRAL DE RAZONES DE VARIANZAS

5.79. De una población distribuida de manera normal se toman dos muestras de tamaño 4 y 8. ¿Existe la probabilidad de que una de las varianzas sea 1.5 veces mayor que la otra, 0.05, entre 0.05 y 0.01 o menos de 0.01?

5.80. Dos empresas, A y B, fabrican bombillas eléctricas, cuyas duraciones están distribuidas normalmente. La duración de las bombillas de A tiene una desviación estándar de 40 horas, mientras que la desviación estándar de las de B es de 50 horas. De A se toma una muestra de 8 bombillas y de B otra de 16 bombillas. Calcule la probabilidad de que la varianza de la primera muestra sea *a*) más del doble de la varianza de la segunda, *b*) más de 1.2 veces la de la segunda.

5.81. Repita el problema 5.80 considerando que las desviaciones estándar de las duraciones son *a*) las dos de 40 horas, *b*) las dos de 50 horas.

DISTRIBUCIÓN DE FRECUENCIAS

5.82. En la tabla 5-16 se muestra una distribución de frecuencias de la duración de 400 cinescopios probados en una empresa. De acuerdo con esta tabla, determine

a) el límite superior de la quinta clase

b) el límite inferior de la octava clase

c) la marca de clase de la séptima clase

Tabla 5-16

Vida media (horas)	Número de cinescopios
300-399	14
400-499	46
500-599	58
600-699	76
700-799	68
800-899	62
900-999	48
1 000-1 099	22
1 100-1 199	6
TOTAL	400

d) límites de clase de la última clase
e) tamaño del intervalo de clase
f) frecuencia de la cuarta clase
g) frecuencia relativa de la sexta clase
h) porcentaje de cinescopios cuya duración no es mayor a 600 horas
i) porcentaje de cinescopios cuya duración es mayor o igual a 900 horas
j) porcentaje de cinescopios cuya duración es de por lo menos 500 horas, pero menor a 1 000 horas.

5.83. Elabore *a*) un histograma, *b*) un polígono de frecuencias que corresponda a la distribución de frecuencias del problema 5.82.

5.84. Con los datos del problema 5.82, elabore *a*) una distribución de frecuencias relativas o porcentuales, *b*) un histograma de frecuencias relativas, *c*) un polígono de frecuencias relativas.

5.85. En el problema 5.82 estime el porcentaje de cinescopios que tienen una duración *a*) menor a 560 horas, *b*) de 970 o más horas, *c*) entre 620 y 890 horas.

5.86. El diámetro interno de las rondanas producidas en una empresa se mide con una exactitud de milésimas. Si las marcas de clase en una distribución de frecuencias de estos diámetros son 0.321, 0.324, 0.327, 0.330, 0.333 y 0.336, encuentre *a*) el tamaño del intervalo de clase, *b*) las fronteras de clase, *c*) los límites de clase.

5.87. En la tabla 5-17 se muestran los diámetros en pulgadas de una muestra de 60 balines. Elabore una distribución de frecuencias con estos diámetros usando los intervalos de clase adecuados.

Tabla 5-17

0.738	0.729	0.743	0.740	0.736	0.741	0.735	0.731	0.726	0.737
0.728	0.737	0.736	0.735	0.724	0.733	0.742	0.736	0.739	0.735
0.745	0.736	0.742	0.740	0728	0.738	0.725	0.733	0.734	0.732
0.733	0.730	0.732	0.730	0.739	0.734	0.738	0.739	0.727	0.735
0.735	0.732	0.735	0.727	0.734	0.732	0.736	0.741	0.736	0.744
0.732	0.737	0.731	0.746	0.735	0.735	0.729	0.734	0.730	0.740

5.88. Con los datos del problema 5.87 elabore *a*) un histograma, *b*) un polígono de frecuencias, *c*) una distribución de frecuencias relativas, *d*) un histograma de frecuencias relativas, *e*) un polígono de frecuencias relativas.

5.89. De acuerdo con los resultados del problema 5.88, determine el porcentaje de balines que tienen diámetros *a*) mayores a 0.732 pulgadas, *b*) no mayores a 0.736 pulgadas, *c*) entre 0.730 y 0.738 pulgadas. Compare los resultados con los obtenidos directamente de los datos en bruto de la tabla 5-17.

5.90. Repita el problema 5.88 con los datos del problema 5.82.

CÁLCULO DE LA MEDIA, DESVIACIÓN ESTÁNDAR Y LOS MOMENTOS MUESTRALES

5.91. Las calificaciones de un estudiante en cinco materias son 85, 76, 93, 82 y 96. Determine la media aritmética de estas calificaciones.

5.92. Un psicólogo mide los tiempos de reacción de un individuo a cierto estímulo y encuentra que éstos son de 0.53, 0.46, 0.50, 0.49, 0.52, 0.53, 0.44 y 0.55 segundos. Determine el tiempo medio de reacción del individuo a este estímulo.

5.93. Un conjunto de números consta de 6 seises, 7 sietes, 8 ochos, 9 nueves y 10 dieces. ¿Cuál es la media aritmética de estos números?

5.94. Las calificaciones que obtuvo un estudiante en tres aspectos de un curso son 71, 78 y 89. *a*) Si los pesos que se dan a estas calificaciones son 2, 4 y 5, respectivamente, ¿cuál es el promedio? *b*) ¿Cuál es el promedio si se usan pesos iguales?

5.95. Las calificaciones promedio de los grupos de tres profesores son 79, 74 y 82; sus grupos son de 32, 25 y 17 estudiantes, respectivamente. Encuentre la calificación promedio de todas las clases.

5.96. El salario medio anual pagado a todos los empleados de una empresa es de $5 000. El salario medio anual pagado a los empleados y a las empleadas de la empresa es $5 200 y $4 200, respectivamente. Estime el porcentaje de empleados y de empleadas de esta empresa.

5.97. En la tabla 5-18 se muestra la distribución de las cargas máximas en toneladas cortas (1 tonelada corta = 2 000 lb) que resisten ciertos cables de una empresa. Determine la carga máxima media con base en *a*) el "método largo", *b*) el método de compilación.

Tabla 5-18

Carga máxima (toneladas cortas)	Número de cables
9-9.7	2
9.8-10.2	5
10.3-10.7	12
10.8-11.2	17
11.3-11.7	14
11.8-12.2	6
12.3-12.7	3
12.8-13.2	1
TOTAL	60

Tabla 5-19

x	462	480	498	516	534	552	570	588	606	624
f	98	75	56	42	30	21	15	11	6	2

5.98. Con base en los datos de la tabla 5-19 estime x usando *a*) el método largo, *b*) el método de compilación.

5.99. En la tabla 5-20 se muestra una distribución de diámetros de remaches. Calcule el diámetro medio.

Tabla 5-20

Diámetro (pulgadas)	Frecuencia
0.7247-0.7249	2
0.7250-0.7252	6
0.7253-0.7255	8
0.7256-0.7258	15
0.7259-0.7261	42
0.7262-0.7264	68
0.7265-0.7267	49
0.7268-0.7270	25
0.7271-0.7273	18
0.7274-0.7276	12
0.7277-0.7279	4
0.7280-0.7282	1
TOTAL	250

5.100. Calcule la media de los datos de la tabla 5-21.

Tabla 5-21

Clase	Frecuencias
10-y menor que 15	3
15-y menor que 20	7
20-y menor que 25	16
25-y menor que 30	12
30-y menor que 35	9
35-y menor que 40	5
40-y menor que 45	2
TOTAL	54

5.101. Encuentre la desviación estándar de los números:
a) 3, 6, 2, 1, 7, 5; *b*) 3.2, 4.6, 2.8, 5.2, 4.4; *c*) 0, 0, 0, 0, 0, 1, 1, 1.

5.102. *a*) Sumando 5 a cada uno de los números del conjunto 3, 6, 2, 1, 7, 5, se obtiene el conjunto 8, 11, 7, 6 12, 10. Demuestre que estos dos conjuntos tienen la misma desviación estándar, pero diferentes medias. ¿Cuál es la relación entre las medias?

b) Si multiplica cada uno de los números 3, 6, 2, 1, 7, 5 por 2 y después les suma 5, obtiene el conjunto 11, 17, 9, 7, 19, 15. ¿Qué relación hay entre las desviaciones estándar y las medias de estos dos conjuntos?

c) ¿Qué propiedades de la media y de la desviación estándar se ilustran mediante los conjuntos de números de los incisos *a*) y *b*)?

5.103. Estime la desviación estándar del conjunto de números de la serie aritmética 4, 10, 16, 22, . . . , 154.

5.104. Calcule la desviación estándar de la distribución de *a*) el problema 5.97, *b*) el problema 5.98.

5.105. Determine *a*) la media, *b*) la desviación estándar de la distribución del problema 5.30 y explique el significado de los resultados que obtenga.

5.106. *a*) Estime la desviación estándar s de los diámetros del problema 5.99. *b*) ¿Qué porcentaje de estos diámetros se encuentra entre $(\bar{x} \pm s)$, $(\bar{x} \pm 2s)$, $(\bar{x} \pm 3s)$? *c*). Compare los porcentajes que obtuvo en el inciso *b*) con los que se deberían esperar teóricamente si la distribución fuera normal y explique las diferencias que observó.

5.107. *a*) Dados los datos del problema 5.28, calcule la media y la desviación estándar.

b) Elabore una distribución de frecuencias con estos datos y determine la desviación estándar.

c) Compare los resultados de los incisos *a*) y *b*).

5.108. Repita el problema 5.107 con los datos del problema 5.87.

5.109. *a*) En un total de n números, la proporción p es de unos y la proporción $q = 1 - p$ es de ceros. Pruebe que la desviación estándar de este conjunto de números es $\sqrt{pq}$. *b*). Aplique el resultado del inciso *a*) al problema 5.101*c*).

5.110. Dado el conjunto de números 4, 7, 5, 9, 8, 3, 6, calcule el *a*) primer, *b*) segundo, *c*) tercer, *d*) cuarto momentos respecto al origen.

5.111. Dado el conjunto de números del problema 5.110, encuentre el *a*) primer, *b*) segundo, *c*) tercer, *d*) cuarto momentos respecto a la media.

5.112. Dado el conjunto de números del problema 5.110, encuentre el *a*) primer, *b*) segundo, *c*) tercer, *d*) cuarto momentos respecto al número 7.

5.113. Con base en los resultados de los problema 5.110 y 5.111, verifique la relación siguiente entre los momentos:

$$a)\ m_2 = m'_2 - m'^2_1,\ b)\ m_3 = m'_3 - 3m'_1m'_2 + 2m'^3_1,\ c)\ m_4 = m'_4 - 4m'_1m'_3 + 6m'^2_1m'_2 - 3m'^4_1.$$

5.114. Dado el conjunto de números de la progresión aritmética 2, 5, 8, 11, 14, 17, calcule los cuatro primeros momentos respecto de la media.

5.115. Si el primer momento respecto del número 2 es igual a 5, ¿cuál es la media?

5.116. Si los primeros cuatro momentos respecto del número 3 de un conjunto de números son -2, 10, -25 y 50, determine los momentos correspondientes alrededor de *a*) la media, *b*) el número 5, *c*) el cero.

5.117. Dado el conjunto de números 0, 0, 0, 1, 1, 1, 1, 1, estime los cuatro primeros momentos respecto de la media.

5.112. Dado el conjunto de números del problema 5.110, encuentre el *a*) primer, *b*) segundo, *c*) tercer, *d*) cuarto momentos respecto al número 7.

5.113. Con base en los resultados de los problema 5.110 y 5.111, verifique la relación siguiente entre los momentos: *a*) $m_2 = m'_2 - m'^2_1$, *b*) $m_3 = m'_3 - 3m'_1m'_2 + 2m'^3_1$, *c*) $m_4 = m'_4 - 4m'_1m'_3 + 6m'^2_1m'_2 - 3m'^4_1$.

5.114. Dado el conjunto de números de la progresión aritmética 2, 5, 8, 11, 14, 17, calcule los cuatro primeros momentos respecto de la media.

5.115. Si el primer momento respecto del número 2 es igual a 5, ¿cuál es la media?

5.116. Si los primeros cuatro momentos respecto del número 3 de un conjunto de números son -2, 10, -25 y 50, determine los momentos correspondientes alrededor de *a*) la media, *b*) el número 5, *c*) el cero.

5.117. Dado el conjunto de números 0, 0, 0, 1, 1, 1, 1, 1, estime los cuatro primeros momentos respecto de la media.

5.118. *a*) Pruebe que $m_5 = m'_5 - 5m'_1m'_4 + 10m'^2_1m'_3 - 10m'^3_1m'_2 + 4m'^5_1$. *b*) Obtenga una fórmula similar para m_6.

5.119. En un total de n números, la proporción p es de unos y la proporción $q = 1 - p$ es de ceros. Calcule *a*) m_1, *b*) m_2, *c*) m_3, *d*) m_4. Compárelos con el problema 5.117.

5.120. Dada la distribución de la tabla 5-22, estime los cuatro primeros momentos respecto de la media.

Tabla 5-22

x	f
12	1
14	4
16	6
18	10
20	7
22	2
TOTAL	30

5.121. Dada la distribución del problema 5.97, calcule los cuatro primeros momentos respecto de la media.

5.122. Dada la distribución del problema 5.100, determine *a*) m_1, *b*) m_2, *c*) m_3, *d*) m_4, *e*) $\bar{x}$, *f*) s, *g*) $\overline{x^2}$, *h*) $\overline{x^3}$, *i*) $\overline{x^4}$, *j*) $\overline{(x+1)^3}$.

5.123. Dada la distribución del problema 5.120, estime el coeficiente *a*) de sesgo, *b*) de curtosis.

5.124. Dada la distribución del problema 5.97, encuentre el coeficiente *a*) de sesgo, *b*) de curtosis. Vea el problema 5.121.

5.125. Los segundos momentos respecto de la media de dos distribuciones son 9 y 16, y los terceros momentos respecto de la media son -8.1 y -12.8, respectivamente. ¿Qué distribución está más sesgada hacia la izquierda?

5.126. Los cuartos momentos respecto de la media de las dos distribuciones del problema 5.125 son 230 y 780, respectivamente. ¿Qué distribución se aproxima más a la distribución normal desde el punto de vista *a*) de lo puntiagudo, *b*) del sesgo?

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

5.49. *a*) 9.0 *b*) 4.47 *c*) 9.0 *d*) 3.16 **5.50.** *a*) 9.0 *b*) 4.47 *c*) 9.0 *d*) 2.58

5.51. *a*) $\mu_{\bar{X}} = 22.40$ onzas, $\sigma_{\bar{X}} = 0.008$ onzas *b*) $\mu_{\bar{X}} = 22.40$ onzas, $\sigma_{\bar{X}}$ es ligeramente menor que 0.008 onzas

5.52. *a*) $\mu_{\bar{X}} = 22.40$ onzas, $\sigma_{\bar{X}} = 0.008$ onzas *b*) $\mu_{\bar{X}} = 22.40$ onzas, $\sigma_{\bar{X}} = 0.0057$ onzas

5.53. *a*) 237 *b*) 2 *c*) ninguno *d*) 24 **5.54.** *a*) 0.4972 *b*) 0.1587 *c*) 0.0918 *d*) 0.9544

5.55. *a*) 0.8164 *b*) 0.0228 *c*) 0.0038 *d*) 1.0000 **5.56.** 0.0026

5.57. *a*) 0.0029 *b*) 0.9596 *c*) 0.1446 **5.58.** *a*) 2 *b*) 996 *c*) 218

5.59. *a*) 0.0179 *b*) 0.8664 *c*) 0.1841 **5.60.** *a*) 6 *b*) 9 *c*) 2 *d*) 12

5.62. *a*) 19 *b*) 125 **5.63.** *a*) 0.0077 *b*) 0.8869 **5.64.** *a*) 0.0028 *b*) 0.9172

5.65. *a*) 0.2150 *b*) 0.0064 *c*) 0.4504 **5.66.** 0.0482 **5.67.** 0.0188 **5.68.** 0.0410

5.70. *a*) 118.79 lb *b*) 0.74 lb **5.71.** 0.0228 **5.72.** *a*) 10.00 *b*) 11.49

5.73. *a*) 40/3 *b*) 28.10 **5.74.** *a*) 0.50 *b*) 0.17 *c*) 0.28 **5.75.** *a*) 0.36 *b*) 0.49

5.80. *a*) entre 0.01 y 0.05 *b*) mayor que 0.05 **5.81.** *a*) mayor que 0.05 *b*) mayor que 0.05

5.82. *a*) 799 *c*) 949.5 *e*) 100 (horas) *g*) $62/400 = 0.155$ o 15.5% *i*) 19.0%
b) 1 000 *d*) 1 099.5, 1 199.5 *f*) 76 *h*) 29.5% *j*) 78.0%

5.85. *a*) 24% *b*) 11% *c*) 46%

5.86. *a*) 0.003 pulgadas *b*) 0.3195, 0.3225, 0.3255, . . . , 0.3375 pulgadas
c) 0.320-0.322, 0.323-0.325, 0.326-0.328, . . . , 0.335-0.337

5.91. 86 **5.92.** 0.50 s **5.93.** 8.25 **5.94.** *a*) 82 *b*) 79 **5.95.** 78 **5.96.** 80%, 20%

5.97. 11.09 toneladas **5.98.** 501.0 **5.99.** 0.72642 pulgadas **5.100.** 26.2

5.101. *a*) 2.16 *b*) 0.90 *c*) 0.484 **5.103.** 45 **5.104.** *a*) 0.733 toneladas *b*) 38.60

5.105. *a*) $\bar{x} = 2.47$ *b*) $s = 1.11$ **5.106.** *a*) 0.000576 pulgadas *b*) 72.1%, 93.3%, 99.76%

5.107. *a*) 146.8 lb, 12.9 lb **5.108.** *a*) 0.7349 pulgadas, 0.00495 pulgadas

5.110. *a*) 6 *b*) 40 *c*) 288 *d*) 2 188 **5.111.** *a*) 0 *b*) 4 *c*) 0 *d*) 25.86

5.112. *a*) −1 *b*) 5 *c*) −91 *d*) 53 **5.114.** 0, 26.25, 0, 1 193.1 **5.115.** 7

5.116. *a*) 0, 6, 19, 42 *b*) −4, 22, −117, 560 *c*) 1, 7, 38, 155

5.117. 0, 0.2344, −0.0586, 0.0696 **5.120.** $m_1 = 0, m_2 = 5.97, m_3 = -3.97, m_4 = 89.22$

5.121. $m_1 = 0, m_2 = 0.53743, m_3 = 0.36206, m_4 = 0.84914$

5.122. *a*) 0 *c*) 92.35 *e*) 26.2 *g*) 739.38 *i*) 706 428
b) 52.95 *d*) 7 158.20 *f*) 7.28 *h*) 22 247 *j*) 24 545

5.123. *a*) −0.2464 *b*) 2.62 **5.124.** *a*) 0.9190 *b*) 2.94

5.125. Primera distribución **5.126.** *a*) segunda *b*) primera **5.127.** *a*) 7.2 *b*) 8.4

5.128. *a*) 106 *b*) 4 **5.129.** 159 **5.130.** *a*) 78.7 *b*) 0.0090

PROBLEMAS APORTADOS

1. Un grupo de 200 estudiantes cuya altura media es de 60.96 pulgadas es dividida en dos grupos, uno teniendo una altura media de 63.4 pulgadas, el otro grupo teniendo una altura de media de 57.3 pulgadas, ¿cuántos estudiantes hay en cada grupo?

2. Para cada grupo de números 6.5, 8.3, 4.7, 9.2, 11.3, 8.5, 9.5, 9.2, determine:

a) La media.
b) La mediana.
c) La desviación media.
d) La desviación estándar.

3. Dada la distribución de frecuencia siguiente:

Clase	Frecuencia
10-30	5
30-50	8
50-70	12
70-90	18
90-110	3
110-130	2

Determine:

a) La media.
b) La mediana.
c) La clase modal.
d) La desviación media.
e) La desviación estándar.

4. Una población de cinco panaderías en un municipio en particular reportan ganancias mensuales de $30 000, $25 000, $15 000, $45 000 y $10 000. Obtenga todas las muestras posibles de tamaño 3 sin reemplazo de las ganancias de estas panaderías. Se pide:

a) La media poblacional.
b) La desviación estándar de la población.
c) La media de la distribución muestral de medias.
d) La desviación estándar de la distribución muestral de medias.
e) ¿Los resultados de los incisos *a*) y *c*) son los mismos? Explique.
f) ¿Los resultados de los incisos *b*) y *d*) son los mismos? Explique.

CAPÍTULO 6

ESTIMADORES INSESGADOS Y EFICIENTES

Como se comentó en el capítulo 5 (vea la página 158), se denomina estadístico a un *estimador insesgado* de un parámetro poblacional si la media o la esperanza del estadístico es igual al parámetro. Entonces, el valor correspondiente del estadístico se llama *estimación insesgada* del parámetro.

EJEMPLO 6.1 La media $\bar{X}$ y la varianza $\hat{S}^2$, como se definen en las páginas 155 y 158, son estimadores insesgados de la media poblacional μ y varianza σ^2, puesto que $E(\bar{X}) = \mu$, $E(\hat{S}^2) = \sigma^2$. Por tanto, los valores $\bar{X}$ y $\hat{s}^2$ se llaman *estimadores insesgados*. Sin embargo, $\hat{S}$ es en realidad un estimador insesgado de σ, puesto que en general $E(\hat{S}) \neq \sigma$.

Si las distribuciones muestrales de dos estadísticos tienen la misma media, el estadístico con una varianza más pequeña se denomina el *estimador más eficiente* de la media. El valor correspondiente del estadístico eficiente se llama *estimación eficiente*. Evidentemente, en la práctica se prefiere tener estimaciones que son tanto eficientes como insesgadas, pero esto no siempre es posible.

EJEMPLO 6.2 En el caso de una población normal, la distribución muestral de la media y de la mediana tienen la misma media, es decir, la media poblacional. Sin embargo, la varianza de la distribución muestral de la media es más pequeña que la que corresponde a la distribución muestral de las medianas. Por tanto, la media proporciona una estimación más eficiente que la mediana. Vea la tabla 5-1, página 160.

ESTIMACIONES PUNTUALES Y POR INTERVALOS: CONFIABILIDAD

La estimación de un parámetro poblacional dada por un solo número se llama *estimación puntual* del parámetro. Una estimación del parámetro poblacional dada por dos números, entre los cuales se puede considerar que está el parámetro, se denomina *intervalo de estimación* del parámetro.

EJEMPLO 6.3 Si se dice que una distancia es de 5.28 pies, se proporciona una estimación puntual. Si, por otro lado, se dice que la distancia es de 5.28 ± 0.03 pies, esto es, la distancia se ubica entre 5.25 y 5.31 pies, se da una estimación de intervalo.

Con frecuencia, una afirmación del error o precisión de una estimación se llama su *confiabilidad.*

ESTIMACIONES DEL INTERVALO DE CONFIANZA DE PARÁMETROS POBLACIONALES

Sean μ_S y σ_S la media y la desviación estándar (el error estándar) de la distribución muestral de un estadístico S. Entonces, si la distribución muestral de S es aproximadamente normal (lo que se ha visto que es verdad para muchos estadísticos si el tamaño de la muestra es $n \geq 30$), se espera encontrar S ubicado entre los intervalos $\mu_S - \sigma_S$ a $\mu_S + \sigma_S$, $\mu_S - 2\sigma_S$ a $\mu_S + 2\sigma_S$ o $\mu_S - 3\sigma_S$ a $\mu_S + 3\sigma_S$ alrededor de 68.27%, 95.45% y 99.73% de las veces, respectivamente.

De igual manera, espera encontrar, o podemos tener *confianza* de encontrar μ_S en los intervalos de $S - \sigma_S$ a $S + \sigma_S$, $S - 2\sigma_S$ a $S + 2\sigma_S$ o $S - 3\sigma_S$ a $S + 3\sigma_S$ alrededor de 68.27%, 95.45% y 99.73% de las veces, respectivamente. En consecuencia, los intervalos respectivos se llaman *intervalos de confianza* de 68.27%, 95.45% y 99.73% para estimar μ_S (esto es, para estimar el parámetro poblacional, en el caso de un S insesgado). Entonces, los números finales de estos intervalos ($S \pm \sigma_S$, $S \pm 2\sigma_S$, $S \pm 3\sigma_S$) se denominan los *límites de confianza* de 68.27%, 95.45% y 99.73%.

De la misma manera, $S \pm 1.96\sigma_S$ y $S \pm 2.58\sigma_S$ son límites de confianza de 95% y 99% (o 0.95 y 0.99) para μ_S. Con frecuencia, al porcentaje de confianza se le llama *nivel de confianza*. Los números 1.96, 2.58, etc., en los límí-

tes de confianza, se denominan *valores críticos* y se denotan por z_c. A partir de los niveles de confianza se pueden encontrar valores críticos y viceversa.

En la tabla 6-1 se presentan los valores de z_c correspondientes a varios niveles de confianza que se usan en la práctica. Para niveles de confianza que no se presentan en la tabla, es posible obtener los valores de z_c de la tabla del área de la curva normal del apéndice C.

Tabla 6-1

Nivel de confianza	99.73%	99%	98%	96%	95.45%	95%	90%	80%	68.27%	50%
z_c	3.00	2.58	2.33	2.05	2.00	1.96	1.645	1.28	1.00	0.6745

En los casos en que un estadístico tiene una distribución muestral que es diferente a la distribución normal (como ji cuadrada, t o F), se tienen que hacer modificaciones adecuadas para obtener los intervalos de confianza.

INTERVALOS DE CONFIANZA PARA MEDIAS

1. **MUESTRAS GRANDES ($n \geq 30$).** Si el estadístico S es la media muestral $\bar{X}$, entonces los límites de confianza 95% y 99% para la estimación de la media poblacional μ están dados por $\bar{X} \pm 1.96\sigma_{\bar{X}}$ y $\bar{X} \pm 2.58\sigma_{\bar{X}}$, respectivamente. De manera más general, los límites de confianza están dados por $\bar{X} \pm z_c\sigma_{\bar{X}}$ donde z_c, que depende del nivel particular de confianza, puede leerse en la tabla anterior. Usando los valores de $\sigma_{\bar{X}}$ que se obtuvieron en el capítulo 5, se puede observar que los límites de confianza de la media poblacional están dados por

$$\bar{X} \pm z_c \frac{\sigma}{\sqrt{n}} \tag{1}$$

en el caso en que la muestra es de una población infinita o si la muestra es con reposición, o devolución, en una población finita y por

$$\bar{X} \pm z_c \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \tag{2}$$

si la muestra es sin reposición de una población finita de tamaño N.

En general, no se conoce la desviación estándar poblacional σ, por lo cual, para obtener los límites de confianza anteriores, se usa el estimador $\hat{S}$ o S.

2. **MUESTRAS PEQUEÑAS ($n < 30$) Y POBLACIÓN NORMAL.** En este caso se usa la distribución t para obtener los niveles de confianza. Por ejemplo, si $-t_{0.975}$ y $t_{0.975}$ son los valores de T para los cuales 2.5% del área se encuentra en la cola de la distribución t, entonces el intervalo de confianza de 95% de T está dado por (vea la página 159),

$$-t_{0.975} < \frac{(\bar{X} - \mu)\sqrt{n}}{\hat{S}} < t_{0.975} \tag{3}$$

por lo que se observa, se puede estimar que μ está en el intervalo

$$\bar{X} - t_{0.975}\frac{\hat{S}}{\sqrt{n}} < \mu < \bar{X} + t_{0.975}\frac{\hat{S}}{\sqrt{n}} \tag{4}$$

con una confianza de 95%. En general, los límites de confianza de las medias poblacionales están dados por

$$\bar{X} \pm t_c \frac{\hat{S}}{\sqrt{n}} \tag{5}$$

el valor t_c se puede leer en el apéndice D.

Una comparación de (5) con (1) indica que para muestras pequeñas se sustituye z_c por t_c. Para $n \geq 30$, z_c y t_c son prácticamente iguales. Cabe mencionar que una ventaja de la teoría de muestras pequeñas (lo cual, evidentemente, también se puede usar para muestras grandes, esto es, muy *exacto*) es que $\hat{S}$ está en (5), por lo cual es posible usar la desviación estándar del muestreo en vez de la desviación estándar poblacional (la que por lo general no se conoce) como en (1).

INTERVALO DE CONFIANZA PARA PROPORCIONES

Suponga que el estadístico S es la proporción de "éxitos" de una muestra de tamaño $n \geq 30$ tomada de una población binomial donde p es la proporción de éxitos (esto es, la probabilidad de éxito). En este caso, los límites de confianza de p están dados por $P \pm z_c\sigma_P$, donde P denota la proporción de éxitos de la muestra de tamaño n. Mediante el uso de los valores de σ_P, que se obtuvieron en el capítulo 5, se advierte que los límites de confianza de la proporción poblacional están dados por

$$P \pm z_c\sqrt{\frac{pq}{n}} = P \pm z_c\sqrt{\frac{p(1-p)}{n}} \tag{6}$$

en el caso en que el muestreo sea en población infinita o si el muestreo es con reposición en una población finita. De manera similar, los límites de confianza son

$$P \pm z_c\sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}} \tag{7}$$

si el muestreo es sin reposición de una población finita de tamaño N. Observe que estos resultados se obtienen de (1) y (2) al sustituir $\bar{X}$ por P y σ por $\sqrt{pq}$.

Para calcular los límites de confianza anteriores, se usa el estimador muestral P para p. Un método más preciso está dado en el problema 6.27.

INTERVALOS DE CONFIANZA PARA DIFERENCIAS Y SUMAS

Si S_1 y S_2 son dos estadísticos muestrales con distribuciones aproximadamente normales, los límites de confianza para las diferencias de los parámetros poblacionales correspondientes a S_1 y S_2 están dados por

$$S_1 - S_2 \pm z_c\sigma_{S_1-S_2} = S_1 - S_2 \pm z_c\sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2} \tag{8}$$

mientras que los límites de confianza para la suma de los parámetros poblacionales están dados por

$$S_1 + S_2 \pm z_c\sigma_{S_1+S_2} = S_1 + S_2 \pm z_c\sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2} \tag{9}$$

siempre y cuando las muestras sean independientes.

Por ejemplo, los límites de confianza para la diferencia entre dos medias poblacionales, en el caso en que las poblaciones son infinitas y tienen desviaciones estándar conocidas σ_1, σ_2, están dados por

$$\bar{X}_1 - \bar{X}_2 \pm z_c\sigma_{\bar{X}_2-\bar{X}_2} = \bar{X}_1 - \bar{X}_2 \pm z_c\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{10}$$

donde $\bar{X}_1$, n_1 y $\bar{X}_2$, n_2 son las medias y tamaños respectivos de las dos muestras tomadas de las poblaciones.

De manera similar, los límites de confianza para la diferencia entre dos proporciones de poblaciones, donde éstas son infinitas, están dados por

$$P_1 - P_2 \pm z_c\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}} \tag{11}$$

donde P_1 y P_2 son las dos proporciones muestrales, y n_1 y n_2 son los tamaños de las dos muestras tomadas de las poblaciones.

INTERVALOS DE CONFIANZA PARA LA VARIANZA DE UNA DISTRIBUCIÓN NORMAL

El hecho de que $nS^2/\sigma^2 = (n-1)\hat{S}^2/\sigma^2$ tiene una distribución ji cuadrada con $n-1$ grados de libertad permite obtener los límites de confianza para σ^2 o para σ. Por ejemplo, si $\chi^2_{0.025}$ y $\chi^2_{0.975}$ son valores de χ^2 para los cuales 2.5% del área está en cada cola de la distribución, entonces un intervalo de confianza de 95% es

$$\chi^2_{0.025} \leq \frac{nS^2}{\sigma^2} \leq \chi^2_{0.975} \tag{12}$$

o de manera equivalente

$$\chi^2_{0.025} \leq \frac{(n-1)\hat{S}^2}{\sigma^2} \leq \chi^2_{0.975} \tag{13}$$

Aquí se puede ver que σ se puede estimar que está en el intervalo

$$\frac{S\sqrt{n}}{\chi_{0.975}} \leq \sigma \leq \frac{S\sqrt{n}}{\chi_{0.025}} \tag{14}$$

o, de manera equivalente,

$$\frac{\hat{S}\sqrt{n-1}}{\chi_{0.975}} \leq \sigma \leq \frac{\hat{S}\sqrt{n-1}}{\chi_{0.025}} \tag{15}$$

con 95% de confianza. De manera similar, es posible encontrar otros intervalos de confianza.

En general, es deseable que el ancho esperado de un intervalo de confianza sea tan pequeño como sea posible. En estadísticos con distribuciones muestrales simétricas, como la distribución normal y t, eso se logra al usar las colas de áreas iguales. No obstante, para distribuciones no simétricas como la de ji cuadrada, puede ser conveniente ajustar las áreas de las colas para obtener el intervalo más pequeño. Este proceso se ilustra en el problema 6.28.

INTERVALOS DE CONFIANZA PARA RAZONES, O COCIENTES, DE VARIANZAS

En el capítulo 5, página 159, se vio que si se toman dos muestras independientes y aleatorias de tamaños m y n con varianzas S_1^2, S_2^2, de dos poblaciones distribuidas de manera normal con varianzas σ_1^2, σ_2^2, respectivamente, entonces la variable aleatoria $\dfrac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2}$ tiene una distribución F con $m - 1$, $n - 1$ grados de libertad. Por ejemplo, si los valores de F se denotan por $F_{0.01}$ y $F_{0.99}$, para los cuales 1% del área reside en cada cola de la distribución F, entonces con 98% de confianza, se tiene

$$F_{0.01} \leq \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \leq F_{0.99} \tag{16}$$

Aquí se puede ver que un intervalo de confianza de 98% para la razón de varianzas σ_1^2/σ_2^2 de las dos poblaciones está dado por

$$\frac{1}{F_{0.99}}\frac{\hat{S}_1^2}{\hat{S}_2^2} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{1}{F_{0.01}}\frac{\hat{S}_1^2}{\hat{S}_2^2} \tag{17}$$

Observe que $F_{0.99}$ se obtiene de una de las tablas del apéndice F. El valor $F_{0.01}$ es el recíproco de $F_{0.99}$ con los grados de libertad del numerador y el denominador invertidos, de acuerdo con el teorema 4-8, página 117.

De manera similar, se encuentra el intervalo de confianza de 90% al usar la tabla apropiada del apéndice F. Éste estaría dado por

$$\frac{1}{F_{0.95}}\frac{\hat{S}_1^2}{\hat{S}_2^2} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{1}{F_{0.05}}\frac{\hat{S}_1^2}{\hat{S}_2^2} \tag{18}$$

ESTIMACIONES DE LA MÁXIMA VEROSIMILITUD

Aunque son valiosos los límites de confianza para estimar un parámetro poblacional, a menudo es conveniente tener una sola estimación o estimación puntual. Para obtener la "mejor" estimación puntual, se utiliza una técnica conocida como *método de la máxima verosimilitud* de Fisher.

Para ilustrar este método, se supone que la población tiene una función de densidad que contiene un parámetro poblacional θ, lo que se estima usando un estadístico en particular. Entonces, la función de densidad se puede denotar por $f(x, \theta)$. Suponiendo que hay n observaciones independientes, $X_1, \ldots, X_n$, la función de densidad conjunta de estas observaciones es

$$L = f(x_1, \theta)f(x_2, \theta) \cdots f(x_n, \theta) \tag{19}$$

lo que se llama *verosimilitud.* La *máxima verosimilitud* se puede obtener al tomar la derivada de L con respecto a θ e igualarla a cero. Para este propósito es conveniente tomar primero sus logaritmos y después derivar la función. De

$$\frac{}{\theta)}\frac{\partial f(x_1, \theta)}{\partial \theta} + \cdots + \frac{1}{f(x_n, \theta)}\frac{\partial f(x_n, \theta)}{\partial \theta} = 0 \tag{20}$$

La solución de esta ecuación, para θ en términos de x_k, se conoce como *estimador de máxima verosimilitud de* θ.

Es posible generalizar el método. En el caso en que haya varios parámetros, se toman las derivadas parciales con respecto a cada uno de ellos, se igualan a cero y se resuelven de manera simultánea las ecuaciones resultantes.

PROBLEMAS RESUELTOS

ESTIMADORES INSESGADOS Y EFICIENTES

6.1. Dar ejemplos de estimadores (o estimaciones) que son *a*) insesgados y eficientes, *b*) insesgados e ineficientes, *c*) sesgados e ineficientes.

Suponemos que la población es normal. Entonces,

a) La media de la muestra $\bar{X}$ y la varianza modificada de la muestra $\hat{S}^2 = \frac{n}{n-1}S^2$ son dos ejemplos.

b) La mediana de la muestra y el estadístico muestral $\frac{1}{2}(Q_1 + Q_3)$, donde Q_1 y Q_3 son los cuartiles superior e inferior de la muestra, son dos ejemplos de este tipo. Ambos estadísticos son estimaciones insesgadas de la media poblacional, puesto que se puede demostrar que la media de sus distribuciones muestrales es la media poblacional. Sin embargo, ambas son ineficientes en comparación con $\bar{X}$.

c) La desviación estándar muestral S, la desviación estándar modificada $\hat{S}$, la desviación media y el rango semi-intercuartil son cuatro ejemplos para evaluar la desviación estándar poblacional σ.

6.2. Un científico registró, en una muestra, las cinco medidas siguientes del diámetro de una esfera, 6.33, 6.37, 6.36, 6.32 y 6.37 cm. Determinar estimaciones insesgadas y eficientes de *a*) la media verdadera, *b*) la varianza verdadera. Se considera que la medida del diámetro está distribuida normalmente.

a) Una estimación insesgada y eficiente de la media verdadera (es decir, la media poblacional) es

$$\bar{x} = \frac{\sum x}{n} = \frac{6.33 + 6.37 + 6.36 + 6.32 + 6.37}{5} = 6.35 \text{ cm}$$

b) Una estimación insesgada y eficiente de la varianza verdadera (es decir, la varianza poblacional) es

$$\hat{s}^2 = \frac{n}{n-1}s^2 = \frac{\sum(x - \bar{x})^2}{n-1}$$

$$= \frac{(6.33 - 6.35)^2 + (6.37 - 6.35)^2 + (6.36 - 6.35)^2 + (6.32 - 6.35)^2 + (6.37 - 6.35)^2}{5 - 1}$$

$$= 0.00055 \text{ cm}^2$$

Observemos que $\hat{s} = \sqrt{0.00055} = 0.023$ es una estimación de la desviación estándar verdadera, pero esta estimación no es insesgada ni eficiente.

6.3. Se supone que las estaturas de 100 estudiantes masculinos de la Universidad XYZ representan una muestra aleatoria de las estaturas de la totalidad de los 1 546 estudiantes masculinos en la universidad. Determinar la estimación insesgada y eficiente de *a*) la media verdadera, *b*) la varianza verdadera.

a) Del problema 5.33:

Estimación insesgada y eficiente de la media verdadera de estatura $= \bar{x} = 67.45$ pulgadas.

b) Del problema 5.38:

Estimación insesgada y eficiente de la varianza verdadera $= \hat{s}^2 = \frac{n}{n-1}s^2 = \frac{100}{99}(8.5275) = 8.6136$

Por tanto, $\hat{s} = \sqrt{8.6136} = 2.93$. Observemos que como n es grande casi no hay diferencia entre s^2 y $\hat{s}^2$ o entre s y $\hat{s}$.

6.4. Dar una estimación insesgada e ineficiente del verdadero diámetro (media) de la esfera del problema 6.2.

La mediana es un ejemplo de una estimación insesgada e ineficiente de la media poblacional. En el caso de las cinco medidas ordenadas de acuerdo con la magnitud, la mediana corresponde a 6.36 cm.

ESTIMACIONES DEL INTERVALO DE CONFIANZA PARA MEDIAS (MUESTRAS GRANDES)

6.5. Encontramos intervalos de confianza de *a*) 95%, *b*) 99% para estimar la estatura media de los estudiantes de la Universidad XYZ del problema 6.3.

a) Los límites de confianza de 95% son $\bar{X} \pm 1.96\sigma/\sqrt{n}$.

Utilizando $\bar{x} = 67.45$ pulgadas y $\hat{s} = 2.93$ pulgadas como una estimación de σ (vea el problema 6.3), los límites de confianza son $67.45 \pm 1.96(2.93/\sqrt{100})$, o 67.45 ± 0.57, pulgadas. En consecuencia, el intervalo de confianza de 95% para la media poblacional μ es de 66.88 a 68.02 pulgadas, lo que se puede denotar mediante $66.88 < \mu < 68.02$.

Por tanto, podemos decir que la probabilidad de que la estatura media poblacional se encuentre entre 66.88 y 68.02 pulgadas es aproximadamente 95% o 0.95. En símbolos, se escribe $P(66.88 < \mu < 68.02) = 0.95$. Esto equivale a afirmar que tenemos 95% *confianza* de que la media poblacional (o media verdadera) esté entre 66.88 y 68.02 pulgadas.

b) Los límites de confianza de 99% son $\bar{X} \pm 2.58\sigma/\sqrt{n}$. Para la muestra dada,

$$\bar{x} \pm 2.58\frac{\hat{s}}{\sqrt{n}} = 67.45 \pm 2.58\frac{2.93}{\sqrt{100}} = 67.45 \pm 0.76 \text{ pulgadas}$$

Por tanto, el intervalo de confianza de 99% para la media poblacional μ es de 66.69 a 68.21 pulgadas, lo que denotamos como $66.69 < \mu < 68.21$.

Para obtener los intervalos de confianza anteriores, supusimos que la población es infinita o tan grande que sería posible considerar las mismas condiciones que en el muestreo con reposición. En el caso de poblaciones finitas donde el muestreo es sin reposición, se debe usar $\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ en vez de $\frac{\sigma}{\sqrt{n}}$. No obstante, es posible considerar el factor $\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{1\,546 - 100}{1\,546 - 1}} = 0.967$ como 1.0, por lo que no es necesario usarlo. Si se usa, los límites de confianza anteriores son los valores 67.45 ± 0.56 y 67.45 ± 0.73 pulgadas respectivamente.

6.6. Las medidas de los diámetros de una muestra aleatoria de 200 cojinetes de bolas fabricados en una semana por una máquina específica indicaron una media de 0.824 pulgadas y una desviación estándar de 0.042 pulgadas. Determinar los límites de confianza de *a*) 95% y *b*) 99% para la media del diámetro de todos los cojinetes de bolas.

Puesto que $n = 200$ es grande, suponemos que $\bar{X}$ es muy cercana a la normal.

a) Los límites de confianza de 95% son

$$\bar{X} \pm 1.96\frac{\sigma}{\sqrt{n}} = \bar{x} \pm 1.96\frac{\hat{s}}{\sqrt{n}} = 0.824 \pm 1.96\frac{0.042}{\sqrt{200}} = 0.824 \pm 0.0058 \text{ pulgadas}$$

o sea 0.824 ± 0.006 pulgadas.

b) Los límites de confianza de 99% son

$$\bar{X} \pm 2.58\frac{\sigma}{\sqrt{n}} = \bar{x} \pm 2.58\frac{\hat{s}}{\sqrt{n}} = 0.824 \pm 2.58\frac{0.042}{\sqrt{200}} = 0.824 \pm 0.0077 \text{ pulgadas}$$

o sea 0.824 ± 0.008 pulgadas.

Observemos que se considera que la desviación estándar indicada es la desviación estándar *modificada* $\hat{s}$. Si la desviación estándar hubiera sido s, se hubiera usado $\hat{s} = \sqrt{n/(n-1)}s = \sqrt{200/199}\,s$, la cual se toma como s para todos los propósitos prácticos. En general, para $n \geq 30$, es posible considerar que s y $\hat{s}$ son prácticamente iguales.

6.7. Calcular los límites de confianza de *a*) 98%, *b*) 90%, *c*) 99.73% de la media del diámetro de los cojinetes de bolas en el problema 6.6.

a) Sea z_c tal que el área bajo la curva normal a la derecha de $z = z_c$ es 1%. Entonces, por simetría, el área a la izquierda de $z = -z_c$ también es 1%, por lo que el área sombreada es 98% del área total (figura 6-1).

Puesto que el área total bajo la curva es uno, el área de $z = 0$ a $z = z_c$ es 0.49; por ello, $z_c = 2.33$. Por tanto, los límites de confianza de 98% son

$$\bar{x} \pm 2.33\frac{\sigma}{\sqrt{n}} = 0.824 \pm 2.33\frac{0.042}{\sqrt{200}} = 0.824 \pm 0.0069 \text{ pulgadas}$$

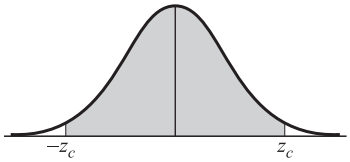


Figura 6-1

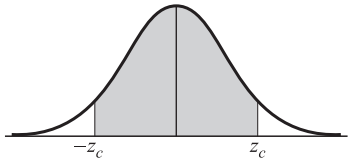


Figura 6-2

b) Se requiere z_c para que el área de $z = 0$ a $z = z_c$ sea 0.45. Por lo que, $z_c = 1.645$ (figura 6-2).

Entonces, los límites de confianza de 90% son

$$\bar{x} \pm 1.645\frac{\sigma}{\sqrt{n}} = 0.824 \pm 1.645\frac{0.042}{\sqrt{200}} = 0.824 \pm 0.0049 \text{ pulgadas}$$

c) Los límites de confianza de 99.73% son

$$\bar{x} \pm 3\frac{\sigma}{\sqrt{n}} = 0.824 \pm 3\frac{0.042}{\sqrt{200}} = 0.824 \pm 0.0089 \text{ pulgadas}$$

6.8. Al medir el tiempo de reacción, un psicólogo estima que la desviación estándar es de 0.05 segundos. ¿De qué tamaño debe ser la muestra de medidas para tener una confianza de *a*) 95% y *b*) 99%, y que el error en la estimación de la media del tiempo de reacción no exceda 0.01 segundos?

a) Los límites de confianza de 95% son $\bar{X} \pm 1.96\sigma/\sqrt{n}$, con el error de la estimación igual a $1.96\sigma/\sqrt{n}$. Considerando $\sigma = s = 0.05$ segundos, se nota que este error es igual a 0.01 segundos si $(1.96)(0.05)/\sqrt{n} = 0.01$, es decir, $\sqrt{n} = (1.96)(0.05)/0.01 = 9.8$, o $n = 96.04$. Por tanto, se tiene una confianza de 95% de que el error en la estimación será menor que 0.01 si n es 97 o mayor.

b) Los límites de confianza de 99% son $\bar{X} \pm 2.58\sigma/\sqrt{n}$. Entonces, $(2.58)(0.05/\sqrt{n} = 0.01$, o bien $n = 166.4$. Por tanto, se tiene una confianza de 99% de que el error en la estimación será menor que 0.01 sólo si n es 167 o mayor.

Observemos que la solución anterior considera una distribución cercana a la normal para $\bar{X}$, lo que se justifica, puesto que la n que se obtiene es grande.

6.9. En una muestra aleatoria de 50 calificaciones de matemáticas, de un total de 200, se obtuvo una media de 75 con una desviación estándar de 10. *a*) ¿Cuáles son los límites de confianza de 95% para la media de las 200 calificaciones? *b*) ¿Con qué grado de confianza se puede afirmar que la media de las 200 calificaciones es 75 ± 1?

a) Puesto que la población no es muy grande, en comparación con el tamaño de la muestra, se ajusta para un *muestreo sin reposición*. Entonces, para 95% los límites de confianza son

$$\bar{X} \pm 1.96\sigma_{\bar{X}} = \bar{X} \pm 1.96\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = 75 \pm 1.96\frac{10}{\sqrt{50}}\sqrt{\frac{200-50}{200-1}} = 75 \pm 2.4$$

b) Es posible representar los límites de confianza por

$$\bar{X} \pm z_c\sigma_{\bar{X}} = \bar{X} \pm z_c\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = 75 \pm z_c\frac{(10)}{\sqrt{50}}\sqrt{\frac{200-50}{200-1}} = 75 \pm 1.23z_c$$

Puesto que esto debe ser igual a 75 ± 1, se tiene que $1.23z_c = 1$, o sea $z_c = 0.81$. El área bajo la curva normal de $z = 0$ a $z = 0.81$ es 0.2910. Por tanto, el grado de confianza que se requiere es $2(0.2919) = 0.582$ o 58.2%.

ESTIMACIONES DEL INTERVALO DE CONFIANZA PARA MEDIAS (MUESTRAS PEQUEÑAS)

6.10. Los valores críticos de 95% (dos colas) para la distribución normal están dados por ± 1.96. ¿Cuáles son los valores correspondientes para la distribución t si el número de grados de libertad es *a*) $\nu = 9$, *b*) $\nu = 20$, *c*) $\nu = 30$, *d*) $\nu = 60$?

En el caso de valores críticos de 95% (dos colas), el área sombreada total en la figura 6-3 debe ser 0.05. Por tanto, el área sombreada en la cola derecha es 0.025 y el valor crítico correspondiente es $t_{0.975}$. Entonces, los valores críticos que buscamos son $\pm t_{0.975}$. Para los valores dados de ν, éstos son *a*) ± 2.26, *b*) ± 2.09, *c*) ± 2.04, *d*) ± 2.00.

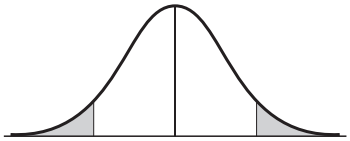


Figura 6-3

6.11. En una muestra de 10 mediciones del diámetro de un esfera se obtuvo una media de $\bar{x} = 4.38$ pulgadas y una desviación estándar de $s = 0.06$ pulgadas. Determinar los límites de confianza para *a*) 95% y *b*) 99% del diámetro real.

a) Los límites de confianza de 95% están dados por $\bar{X} \pm t_{0.975}(S/\sqrt{n-1})$.

Puesto que $\nu = n - 1 = 10 - 1 = 9$, se tiene que $t_{0.975} = 2.26$ [vea también el problema 6.10*a*)]. Entonces, usando $\bar{x} = 4.38$ y $s = 0.06$, los límites de confianza de 95% son

$$4.38 \pm 2.26\frac{0.06}{\sqrt{10-1}} = 4.38 \pm 0.0452 \text{ pulgadas}$$

Por tanto, obtenemos una confianza de 95% de que la media real se encuentra entre $4.38 - 0.045 = 4.335$ pulgadas y $4.38 + 0.045 = 4.425$ pulgadas.

b) Para $\nu = 9$, $t_{0.995} = 3.25$. Por tanto, los límites de confianza de 99% son

$$\bar{X} \pm t_{0.995}(S/\sqrt{n-1}) = 4.38 \pm 3.25(0.06/\sqrt{10-1}) = 4.38 \pm 0.0650 \text{ pulgadas}$$

y el intervalo de confianza de 99% es de 4.315 a 4.445 pulgadas.

6.12. *a*) Resolver el problema 6.11 bajo el supuesto de que son válidos los métodos de la teoría de muestras grandes. *b*) Comparar los resultados de los dos métodos.

a) Con base en la teoría de muestras grandes, los límites de confianza de 95% son

$$\bar{X} \pm 1.96\frac{\sigma}{\sqrt{n}} = 4.38 \pm 1.96\frac{0.06}{\sqrt{10}} = 4.38 \pm 0.037 \text{ pulgadas}$$

donde hemos usado la desviación estándar de la muestra, 0.06, como una estimación de σ. De manera similar, los límites de confianza de 99% son $4.38 \pm (2.58)(0.06)/\sqrt{10} = 4.38 \pm 0.049$ pulgadas.

b) En cada caso, los intervalos de confianza cuando se usan los métodos para muestras pequeñas o exactas son más anchos que los que se obtienen con los métodos para muestras grandes. Esto era previsible, puesto que se dispone de menos precisión con muestras pequeñas que con las grandes.

ESTIMACIÓN DEL INTERVALO DE CONFIANZA PARA PROPORCIONES

6.13. Una muestra de una encuesta de 100 votantes, elegida de manera aleatoria a partir de todos los votantes en un distrito dado, indica que 55% de ellos estaban a favor de cierto candidato. Calcular los límites de confianza para *a*) 95%, *b*) 99% y *c*) 99.73% de la proporción de los votantes que están a favor de este candidato.

a) Los límites de confianza de 95% para la población p son

$$P \pm 1.96\sigma_P = P \pm 1.96\sqrt{\frac{p(1-p)}{n}} = 0.55 \pm 1.96\sqrt{\frac{(0.55)(0.45)}{100}} = 0.55 \pm 0.10$$

donde usamos la proporción muestral 0.55 para estimar p.

b) Los límites de confianza de 99% para p, son: $0.55 \pm 2.58\sqrt{(0.55)(0.45)/100} = 0.55 \pm 0.13$.

c) Los límites de confianza de 99.73% para p son $0.55 \pm 3\sqrt{(0.55)(0.45)/100} = 0.55 \pm 0.15$.

Vea el problema 6.27 donde encontrará el método exacto para resolver este problema.

6.14. ¿De qué tamaño debe ser la muestra de votantes del problema 6.13 para que exista una confianza de 95% de que el candidato será seleccionado?

El candidato se elige si $p > 0.50$, y para tener la confianza de 95% de que va a ser electo, se requiere que prob. $(p > 0.50) = 0.95$. Puesto que $(P - p)/\sqrt{p(1 - p)/n}$ es asintóticamente normal,

$$\text{Prob.}\left(\frac{P - p}{\sqrt{p(1 - p)/n}} < \beta\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\beta} e^{-u^2/2}\,du$$

o bien,

$$\text{Prob.}\,(p > P - \beta\sqrt{p(1 - p)/n}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\beta} e^{-u^2/2}\,du$$

Una comparación con prob.$(p > 0.50) = 0.95$, utilizando el apéndice C, muestra que

$$P - \beta\sqrt{p(1 - p)/n} = 0.50 \qquad \text{donde} \qquad \beta = 1.645$$

Entonces, usando $P = 0.55$ y la estimación $p = 0.55$ del problema 6.13, tenemos

$$0.55 - 1.645\sqrt{(0.55)(0.45)/n} = 0.50 \qquad \text{o sea} \qquad n = 271$$

6.15. En 40 lanzamientos de una moneda se obtuvieron 24 caras. Determinar los límites de confianza para *a*) 95% y *b*) 99.73% de la proporción de caras que se obtendría en un número ilimitado de lanzamientos de la moneda.

a) En el nivel de 95%, $z_c = 1.96$. Al sustituir los valores $P = 24/40 = 0.6$ y $n = 40$ en la fórmula $p = P \pm z_c\sqrt{P(1 - P)/n}$, encontramos que $p = 0.60 \pm 0.15$, que se encuentra en el intervalo de 0.45 a 0.75.

b) En el nivel de 99.73%, $z_c = 3$. Usando la fórmula $p = P \pm z_c\sqrt{P(1 - P)/n}$, encontramos que $p = 0.60 \pm 0.23$, lo que da el intervalo de 0.37 a 0.83.

La fórmula más exacta del problema 6.27 da 95% de confianza en el intervalo de 0.45 a 0.74 y 99.73% de confianza en el intervalo de 0.37 a 0.79.

INTERVALOS DE CONFIANZA PARA DIFERENCIAS Y SUMAS

6.16. Una muestra de 150 focos de marca *A* mostró una vida media de 1 400 horas y una desviación estándar de 120 horas. Una muestra de 200 focos de marca *B* mostró una vida media de 1 200 horas y una desviación estándar de 80 horas. Determinar los límites de confianza para *a*) 95% y *b*) 99% de las vidas medias de las poblaciones de marcas *A* y *B*.

Los límites de confianza de la diferencia de medias de las marcas *A* y *B* están dados por

$$\bar{X}_A - \bar{X}_B \pm z_c\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}$$

a) Los límites de confianza de 95% son $1\,400 - 1\,200 \pm 1.96\sqrt{(120)^2/150 + (80)^2/100} = 200 \pm 24.8$.

Por tanto, se tiene una confianza de 95% de que la diferencia de las medias poblacionales está entre 175 y 225 horas.

b) Los límites de confianza de 99% son $1\,400 - 1\,200 \pm 2.58\sqrt{(120)^2/150 + (80)^2/100} = 200 \pm 32.6$.

Por tanto, se tiene una confianza de 99% de que la diferencia de las medias poblacionales está entre 167 y 233 horas.

6.17. En una muestra aleatoria de 400 adultos y 600 adolescentes que vieron cierto programa de televisión, 100 adultos y 300 adolescentes indicaron que les gustó. Determinar los límites de confianza para *a*) 95% y *b*) 99% de la diferencia de proporciones de todos los adultos y todos los adolescentes que vieron el programa y que les gustó.

Los límites de confianza para la diferencia de proporciones de los dos grupos están dados por

$$P_1 - P_2 \pm z_c\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}$$

donde los subíndices 1 y 2 se refieren a adolescentes y adultos, respectivamente, y $Q_1 = 1 - P_1$, $Q_2 = 1 - P_2$. Aquí, $P_1 = 300/600 = 0.50$ y $P_2 = 100/400 = 0.25$ son, respectivamente, la proporción de adolescentes y adultos a los que les gustó el programa.

a) Los límites de confianza de 95%: $0.50 - 0.25 \pm 1.96\sqrt{(0.50)(0.50)/600 + (0.25)(0.75)/400} = 0.25 \pm 0.06$.

Por tanto, se tiene una confianza de 95% de que la diferencia verdadera en proporciones se encuentra entre 0.19 y 0.31.

b) Los límites de confianza de 99%: $0.50 - 0.25 \pm 2.58\sqrt{(0.50)(0.50)/600 + (0.25)(0.75)/400} = 0.25 \pm 0.08$.

Por tanto, se tiene una confianza de 99% de que la diferencia verdadera en proporciones se encuentra entre 0.17 y 0.33.

6.18. La fuerza electromotriz (fem) de baterías producidas por una compañía se distribuye normalmente con una media de 45.1 voltios y una desviación estándar de 0.04 voltios. Si cuatro de estas baterías se conectan en serie, encuentre los límites de confianza para *a*) 95%, *b*) 99%, *c*) 99.73% y *d*) 50% de la fuerza electromotriz total.

Si E_1, E_2, E_3 y E_4 representan las *fem* de las cuatro baterías, se tiene

$$\mu_{E_1+E_2+E_3+E_4} = \mu_{E_1} + \mu_{E_2} + \mu_{E_3} + \mu_{E_4} \qquad \text{y} \qquad \sigma_{E_1+E_2+E_3+E_4} = \sqrt{\sigma_{E_1}^2 + \sigma_{E_2}^2 + \sigma_{E_3}^2 + \sigma_{E_4}^2}$$

Entonces, puesto que $\mu_{E_1} = \mu_{E_2} = \mu_{E_3} = \mu_{E_4} = 45.1$ voltios y $\sigma_{E_1} = \sigma_{E_2} = \sigma_{E_3} = \sigma_{E_4} = 0.04$ voltios,

$$\mu_{E_1+E_2+E_3+E_4} = 4(45.1) = 180.4 \qquad \text{y} \qquad \sigma_{E_1+E_2+E_3+E_4} = \sqrt{4(0.04)^2} = 0.08$$

a) Los límites de confianza para 95% son $180.4 \pm 1.96(0.08) = 180.4 \pm 0.16$ voltios.

b) Los límites de confianza para 99% son $180.4 \pm 2.58(0.08) = 180.4 \pm 0.21$ voltios.

c) los límites de confianza para 99.73% son $180.4 \pm 3(0.08) = 180.4 \pm 0.24$ voltios.

d) los límites de confianza para 50% son $180.4 \pm 0.6745(0.08) = 180.4 \pm 0.054$ voltios.

El valor de 0.054 voltios se llama *error probable.*

INTERVALOS DE CONFIANZA PARA VARIANZAS

6.19. La desviación estándar de la vida media de una muestra de 200 focos eléctricos se calculó en 100 horas. Encontrar los límites de confianza para *a*) 95% y *b*) 99% de la desviación estándar de todos los focos eléctricos de este tipo.

En este caso, aplicamos la teoría de muestras grandes. Por tanto (vea la tabla 5-1, página 160), los límites de confianza de la desviación estándar poblacional σ están dados por $S \pm z_c\sigma/\sqrt{2n}$, donde z_c indica el nivel de confianza. Se usa la desviación estándar muestral para estimar σ.

a) Los límites de confianza para 95% son $100 \pm 1.96(100)/\sqrt{400} = 100 \pm 9.8$.

Por tanto, se tiene una confianza de 95% de que la desviación estándar poblacional está entre 90.2 y 109.8 horas.

b) Los límites de confianza para 99% son $100 \pm 2.58(100)/\sqrt{400} = 100 \pm 12.9$.

Por tanto, se tiene una confianza de 99% de que la desviación estándar poblacional se ubica entre 87.1 y 112.9 horas.

6.20. ¿De qué tamaño debe ser la muestra de focos en el problema 6.19 para tener una confianza de 99.73% de que la desviación estándar poblacional verdadera no difiera de la desviación estándar muestral más de *a*) 5% y *b*) 10%?

Como en el problema 6.19, los límites de confianza de σ para 99.73% son $S \pm 3\sigma/\sqrt{2n} = s \pm 3s/\sqrt{2n}$, usando s como una estimación de σ. Entonces, el porcentaje de error en la desviación estándar es

$$\frac{3s/\sqrt{2n}}{s} = \frac{300}{\sqrt{2n}}\,\%$$

a) Si $300/\sqrt{2n} = 5$, entonces $n = 1\,800$. Por tanto, el tamaño de la muestra debe ser de 1 800 focos o más.

b) Si $300/\sqrt{2n} = 10$, entonces $n = 450$. Por tanto, el tamaño de la muestra debe ser de 450 focos o más.

6.21. La desviación estándar de las estaturas de 16 estudiantes hombres elegidos de manera aleatoria en una escuela con 1 000 estudiantes hombres es de 2.40 pulgadas. Determinar los límites de confianza para *a*) 95% y *b*) 99%, de la desviación estándar para todos los estudiantes hombres en la escuela. Se supone que la estatura está normalmente distribuida.

a) Los límites de confianza para 95% están dados por $S\sqrt{n}/\chi_{0.975}$ y $S\sqrt{n}/\chi_{0.025}$.

Para $\nu = 16 - 1 = 15$ grados de libertad, $\chi^2_{0.975} = 27.5$ o $\chi_{0.975} = 5.24$ y $\chi^2_{0.025} = 6.26$ o $\chi_{0.025} = 2.50$.

Entonces, los límites de confianza para 95% son $2.40\sqrt{16}/5.24$ y $2.40\sqrt{16}/2.50$, es decir, 1.83 y 3.84 pulgadas. Por tanto, se tiene una confianza de 95% de que la desviación estándar poblacional esté entre 1.83 y 3.84 pulgadas.

b) Los límites de confianza para 99% están dados por $S\sqrt{n}/\chi_{0.995}$ y $S\sqrt{n}/\chi_{0.005}$.

Para $\nu = 16 - 1 = 15$ grados de libertad, $\chi^2_{0.995} = 32.8$ o $\chi^2_{0.995} = 5.73$ y $\chi^2_{0.005} = 4.60$ o $\chi^2_{0.005} = 21.4$.

Por tanto, los límites de confianza para 99% son $2.40\sqrt{16}/5.73$ y $2.40\sqrt{16}/2.14$, esto es, 1.68 y 4.49 pulgadas. Entonces, tenemos una confianza de 99% de que la desviación estándar poblacional está entre 1.68 y 4.49 pulgadas.

6.22. Resolver el problema 6.19 usando la teoría de muestras pequeñas o exactas.

a) Los límites de confianza para 95% están dados por $S\sqrt{n}/\chi_{0.975}$ y $S\sqrt{n}/\chi_{0.025}$.

Para $\nu = 200 - 1 = 199$ grados de libertad, se encuentra como en el problema 4.41, página 136, que

$$\chi^2_{0.975} = \frac{1}{2}(z_{0.975} + \sqrt{2(199) - 1})^2 = \frac{1}{2}(1.96 + 19.92)^2 = 239$$

$$\chi^2_{0.025} = \frac{1}{2}(z_{0.025} + \sqrt{2(199) - 1})^2 = \frac{1}{2}(-1.96 + 19.92)^2 = 161$$

para lo cual $\chi_{0.975} = 15.5$ y $\chi_{0.025} = 12.7$.

Entonces, los límites de confianza para 95% son $100\sqrt{200}/15.5 = 91.2$ y $100\sqrt{200}/12.7 = 111.3$ horas, respectivamente. Por tanto, tenemos una confianza de 95% de que la desviación estándar poblacional está entre 91.2 y 111.3 horas.

Lo anterior puede compararse con el resultado del problema 6.19*a*).

b) Los límites de confianza para 99% están dados por $S\sqrt{n}/\chi_{0.995}$ y $S\sqrt{n}/\chi_{0.005}$.

Para $\nu = 200 - 1 = 199$ grados de libertad,

$$\chi^2_{0.995} = \frac{1}{2}(z_{0.995} + \sqrt{2(199) - 1})^2 = \frac{1}{2}(2.58 + 19.92)^2 = 253$$

$$\chi^2_{0.005} = \frac{1}{2}(z_{0.005} + \sqrt{2(199) - 1})^2 = \frac{1}{2}(-2.58 + 19.92)^2 = 150$$

para lo cual $\chi_{0.995} = 15.9$ y $\chi_{0.005} = 12.2$.

Por ello, los límites de confianza para 99% son $100\sqrt{200}/15.9 = 88.9$ y $100\sqrt{200}/12.2 = 115.9$ horas, respectivamente. Por tanto, se tiene una confianza de 99% de que la desviación estándar poblacional está entre 88.9 y 115.9 horas.

Este resultado debe compararse con el del problema 6.19*b*).

INTERVALOS DE CONFIANZA PARA RAZÓN, O COCIENTES, DE VARIANZA

6.23. Se extraen de manera aleatoria dos muestras de tamaño 16 y 10, respectivamente, de dos poblaciones normales. Si se encuentra que sus varianzas son de 24 y 18, respectivamente, encontrar los límites de confianza para *a*) 98% y *b*) 90% para la razón, o cociente, de las varianzas.

a) Tenemos $m = 16$, $n = 10$, $S_1^2 = 20$, $S_2^2 = 18$, de manera que

$$\hat{s}_1^2 = \frac{m}{m-1}s_1^2 = \left(\frac{16}{15}\right)(24) = 25.2$$

$$\hat{s}_2^2 = \frac{n}{n-1}s_2^2 = \left(\frac{10}{9}\right)(18) = 20.0$$

Del problema 4.47*b*), página 139, se tiene que $F_{0.99} = 4.96$ para $\nu_1 = 16 - 1 = 15$ y $\nu_2 = 10 - 1 = 9$ grados de libertad. También, del problema 4.47*d*), tenemos para $\nu_1 = 15$ y $\nu_2 = 9$ grados de libertad, $F_{0.01} = 1/3.89$, así que $1/F_{0.01} = 3.89$. Entonces, para el intervalo de confianza de 98% se obtiene

$$\left(\frac{1}{4.96}\right)\left(\frac{25.2}{20.0}\right) \le \frac{\sigma_1^2}{\sigma_2^2} \le (3.89)\left(\frac{25.2}{20.0}\right)$$

utilizando (17), página 198, o sea $\quad 0.283 \le \dfrac{\sigma_1^2}{\sigma_2^2} \le 4.90$

b) De igual manera que en *a*), del apéndice F, encontramos que $F_{0.95} = 3.01$ y $F_{0.05} = 1/2.59$. Por tanto, el intervalo de confianza de 90% es

$$\frac{1}{3.01}\left(\frac{25.2}{20.0}\right) \le \frac{\sigma_1^2}{\sigma_2^2} \le (2.59)\left(\frac{25.2}{20.0}\right)$$

o sea, $$0.4186 \le \frac{\sigma_1^2}{\sigma_2^2} \le 3.263$$

Vemos que el intervalo de confianza de 90% es mucho más pequeño que el de 98%, como se esperaba.

6.24. Determinar los límites de confianza para *a*) 98% y *b*) 90% del cociente de las desviaciones estándar del problema 6.23.

Al tomar las raíces cuadradas de las desigualdades del problema 6.23, tenemos que los límites de confianza de 98% y 90% son

a) $$0.53 \le \frac{\sigma_1}{\sigma_2} \le 2.21$$

b) $$0.65 \le \frac{\sigma_1}{\sigma_2} \le 1.81$$

ESTIMACIONES DE LA MÁXIMA VEROSIMILITUD

6.25. Se hacen n observaciones, $X_1, \ldots, X_n$ a partir de una población normalmente distribuida de la que se conoce la varianza pero no la media. Encontrar la estimación de máxima verosimilitud de la media.

Puesto que $$f(x_k, \mu) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x_k-\mu)^2/2\sigma^2}$$

tenemos

(1) $$L = f(x_1, \mu)\cdots f(x_n, \mu) = (2\pi\sigma^2)^{-n/2}e^{-\sum(x_k-\mu)^2/2\sigma^2}$$

Por tanto,

(2) $$\ln L = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum(x_k - \mu)^2$$

Tomando la derivada parcial con respecto a μ se obtiene

(3) $$\frac{1}{L}\frac{\partial L}{\partial \mu} = \frac{1}{\sigma^2}\sum(x_k - \mu)$$

Haciendo $\partial L/\partial\mu = 0$ resulta

(4) $$\sum(x_k - \mu) = 0 \qquad \text{esto es,} \qquad \sum x_k - n\mu = 0$$

o bien,

(5) $$\mu = \frac{\sum x_k}{n}$$

Por tanto, la estimación de máxima verosimilitud es la media de la muestra.

6.26. Si en el problema 6.25 se conoce la media pero no la varianza, encontrar la estimación de máxima verosimilitud de la varianza.

Si se escribe $f(x_k, \sigma^2)$ en vez de $f(x_k, \mu)$, se aplica todo lo que se hizo en el problema 6.25 hasta la ecuación (2). Entonces, tomando la derivada parcial con respecto a σ^2, tenemos

$$\frac{1}{L}\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(x_k - \mu)^2$$

Haciendo $\partial L/\partial\sigma^2 = 0$, encontramos

$$\sigma^2 = \frac{\sum(x_k - \mu)^2}{n}$$

PROBLEMAS DIVERSOS

6.27. *a*) Si P es la proporción observada de éxitos en una muestra de tamaño n, demostrar que los límites de confianza para estimar la proporción de población de éxitos p en el nivel de confianza determinado por z_c están dados por

$$\frac{P + \dfrac{z_c^2}{2n} \pm z_c\sqrt{\dfrac{P(1-P)}{n} + \dfrac{z_c^2}{4n^2}}}{1 + \dfrac{z_c^2}{n}}$$

b) Utilizar la fórmula deducida en *a*) para obtener los límites de confianza de 99.73% del problema 6.13.

c) Demostrar que para n grande la fórmula de *a*) se reduce a $P \pm z_c\sqrt{P(1-P)/n}$, como usamos en el problema 6.13.

a) La proporción de muestra P en unidades estándar es $\dfrac{P-p}{\sigma_P} = \dfrac{P-p}{\sqrt{p(1-p)/n}}$.

Los valores más grandes y más pequeños de esta variable estandarizada son $\pm z_c$, donde z_c determina el nivel de confianza. Por tanto, en estos valores extremos tenemos

$$P - p = \pm z_c\sqrt{\frac{p(1-p)}{n}}$$

Elevando al cuadrado ambos lados,

$$P^2 - 2pP + p^2 = z_c^2\frac{p(1-p)}{n}$$

Multiplicando ambos lados por n y simplificando, encontramos

$$\left(n + z_c^2\right)p^2 - \left(2nP + z_c^2\right)p + nP^2 = 0$$

Si $a = n + z_c^2$, $b = -\left(2nP + z_c^2\right)$ y $c = nP^2$, esta ecuación se convierte en $ap^2 + bp + c = 0$, cuya solución para p está dada por la fórmula cuadrática como

$$p = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{2nP + z_c^2 \pm \sqrt{\left(2nP + z_c^2\right)^2 - 4\left(n + z_c^2\right)(nP^2)}}{2\left(n + z_c^2\right)}$$
$$= \frac{2nP + z_c^2 \pm z_c\sqrt{4nP(1-P) + z_c^2}}{2\left(n + z_c^2\right)}$$

Dividiendo el numerador y denominador entre $2n$, esta expresión se transforma en

$$p = \frac{P + \frac{z_c^2}{2n} \pm z_c\sqrt{\frac{P(1-P)}{n} + \frac{z_c^2}{4n^2}}}{1 + \frac{z_c^2}{n}}$$

b) Para límites de confianza de 99.73%, $z_c = 3$. Entonces, si utilizamos $P = 0.55$ y $n = 100$ en la fórmula deducida en *a*), encontramos $p = 0.40$ y 0.69, lo que concuerda con el problema 6.13*c*).

c) Si n es grande, entonces $z_c^2/2n$, $z_c^2/4n^2$ y z_c^2/n son todos muy pequeños y es posible sustituirlos por cero, por lo que conseguimos el resultado buscado.

6.28. ¿Es posible obtener un intervalo de confianza de 95% de una desviación estándar poblacional cuyo ancho esperado sea más pequeño que el del problema 6.22*a*)?

Los límites de confianza para 95% de la desviación estándar poblacional del problema 6.22*a*) los obtuvimos al elegir valores críticos de χ^2, de tal manera que el área en cada cola fue de 2.5%. Es posible determinar otros límites de confianza de 95% al elegir valores críticos de χ^2 para los cuales la suma de las áreas en las colas es 5% o 0.05, pero tales áreas en cada cola no son iguales.

En la tabla 6-2 se dan algunos valores críticos de este tipo que se han obtenido, y se muestran los intervalos de confianza para 95%.

Tabla 6-2

Valores críticos	Intervalo de confianza de 95%	Ancho
$\chi_{0.01} = 12.44$, $\chi_{0.96} = 15.32$	92.3 a 113.7	21.4
$\chi_{0.02} = 12.64$, $\chi_{0.97} = 15.42$	91.7 a 111.9	20.2
$\chi_{0.03} = 12.76$, $\chi_{0.98} = 15.54$	91.0 a 110.8	19.8
$\chi_{0.04} = 12.85$, $\chi_{0.99} = 15.73$	89.9 a 110.0	20.1

En esta tabla observamos que un intervalo de 95% y ancho solamente de 19.8, va de 91.0 a 110.8.

Es posible encontrar un intervalo con un ancho aún más pequeño si se continúa el mismo método de aproximación, mediante valores críticos como $\chi_{0.031}$ y $\chi_{0.981}$, $\chi_{0.032}$ y $\chi_{0.982}$, etcétera.

A pesar de eso, en general el decrecimiento del intervalo que se obtiene de esta manera es muchas veces insignificante y no vale la pena trabajar en ello.

PROBLEMAS SUPLEMENTARIOS

ESTIMACIONES INSESGADAS Y EFICIENTES

6.29. Se determinan las siguientes mediciones de una muestra de pesos en libras: 8.3, 10.6, 9.7, 8.8, 10.2 y 9.4. Determine las estimaciones insesgadas y eficientes de *a*) la media poblacional y *b*) la varianza poblacional. *c*) Compare la desviación estándar muestral con la desviación estándar poblacional estimada.

6.30. Una muestra de 10 tubos de televisores que produce una compañía indican un tiempo de vida media de 1 200 horas y una desviación estándar de 100 horas. Estime *a*) la media y *b*) la desviación estándar poblacional de todos los tubos de televisión que produce esta compañía.

6.31. *a*) Resuelva el problema 6.30 si se obtienen los mismos resultados para 30, 50 y 100 tubos de televisión. *b*) ¿Qué se puede concluir acerca de la relación entre las desviaciones estándar muestrales y las estimaciones de desviaciones estándar poblacionales de diferentes tamaños de muestras?

ESTIMACIONES DEL INTERVALO DE CONFIANZA PARA MEDIAS (MUESTRAS GRANDES)

6.32. La media y desviación estándar de las máximas cargas soportadas por 60 cables (vea el problema 5.98) son 11.09 toneladas y 0.73 toneladas, respectivamente. Determine los límites de confianza para *a*) 95% y *b*) 99% de la media de las cargas máximas de todos los cables producidos por la compañía.

6.33. La media y desviación estándar de los diámetros de una muestra de 250 cabezas de remaches que fabrica una compañía son 0.72642 y 0.00058 pulgadas, respectivamente (vea el problema 5.99). Calcule los límites de confianza para *a*) 99% , *b*) 98%, *c*) 95%, *d*) 90% de la media del diámetro de todas las cabezas de remache que produce la compañía.

6.34. Determine los límites de confianza de *a*) 50% y *b*) el error probable para la media del diámetro de las cabezas de los remaches en el problema 6.33.

6.35. Si la desviación estándar del tiempo de vida de tubos de televisión se estima en 100 horas, ¿de qué tamaño debe ser la muestra para tener una confianza para *a*) 95%, *b*) 90%, *c*) 99% y *d*) 99.73% de que el error en el tiempo de vida media estimado no exceda de 20 horas?

6.36. ¿Cuáles son los tamaños de muestra del problema 6.35 si el error en el tiempo de vida media estimado no debe exceder de 10 horas?

ESTIMACIONES DEL INTERVALO DE CONFIANZA PARA MEDIAS (MUESTRAS PEQUEÑAS)

6.37. Una muestra de 12 mediciones de la fuerza necesaria para romper hilos de algodón se obtuvo una media de 7.38 onzas y una desviación estándar de 1.24 onzas. Encuentre los límites de confianza para *a*) 95% y *b*) 99% de la media real de la fuerza de rompimiento.

6.38. Resuelva el problema 6.37 bajo el supuesto de que son aplicables los métodos de la teoría de muestras grandes, y compare los resultados obtenidos.

6.39. Se registraron las siguientes cinco mediciones del tiempo de reacción de un individuo a cierto estímulo: 0.28, 0.30, 0.27, 0.33 y 0.31 segundos. Calcule los límites de confianza para *a*) 95% y *b*) 99% de la media real del tiempo de reacción.

ESTIMACIONES DE INTERVALOS DE CONFIANZA PARA PROPORCIONES

6.40. Una urna contiene canicas rojas y blancas en una proporción no conocida. Una muestra aleatoria de 60 canicas seleccionadas con reemplazo mostró que 70% eran rojas. Determine los límites de confianza para *a*) 95%, *b*) 99% y *c*) 99.73% de la proporción real de canicas rojas en la urna. Presente los resultados usando tanto la fórmula aproximada como la fórmula más exacta del problema 6.27.

6.41. ¿De qué tamaño debe tomarse la muestra del problema 6.40 para lograr una confianza de *a*) 95%, *b*) 99% y *c*) 99.73% de que las proporciones real y muestral no tengan una diferencia de más de 5%?

6.42. Se piensa que una elección entre dos candidatos será muy cerrada. Explicar por medio de un ejemplo, enunciando todos los supuestos, cómo se determinaría el menor número de votantes necesarios para encuestarlos acerca de la decisión a favor de cualquiera de los candidatos con una confianza para *a*) 80%, *b*) 95% y *c*) 99%.

INTERVALOS DE CONFIANZA PARA DIFERENCIAS Y SUMAS

6.43. De dos grupos similares de pacientes, *A* y *B*, que constan de 50 y 100 individuos, respectivamente, al primero se le dio un tipo nuevo de somnífero y al segundo un tipo convencional. En los pacientes del grupo *A*, la media del número de horas de sueño fue de 7.82 con una desviación estándar de 0.24 horas. En los pacientes del grupo *B*, la media del número de horas de sueño correspondió a 6.75 con una desviación estándar de 0.30 horas. Determine los límites de confianza de *a*) 95% y *b*) 99% de la diferencia entre el número medio de horas de sueño inducido por los dos tipos de somníferos.

6.44. Una muestra de 200 pernos de una máquina indicó que 15 fueron defectuosos, mientras que una muestra de 100 de otra máquina mostró que 12 lo fueron. Calcule los límites de confianza de *a*) 95%, *b*) 99% y *c*) 99.73% para la diferencia entre proporciones de pernos defectuosos de las dos máquinas. Explique los resultados que obtenga.

6.45. Una compañía fabrica cojinetes de bolas que tienen un peso medio de 0.638 onzas y una desviación estándar de 0.012 onzas. Encuentre los límites de confianza de *a*) 95% y *b*) 99% de los pesos de lotes que constan de 100 cojinetes de bolas cada uno.

INTERVALOS DE CONFIANZA PARA VARIANZAS O DESVIACIONES ESTÁNDAR

6.46. La desviación estándar de las resistencias a la ruptura de 100 cables probados por una compañía fue de 1 800 libras. Calcule los límites de confianza de *a*) 95%, *b*) 99% y *c*) 99.73% para la desviación estándar de todos los cables producidos por la compañía.

6.47. ¿De qué tamaño tiene que ser una muestra para que tenga una confianza de *a*) 95%, *b*) 99% y *c*) 99.73% de que una desviación estándar poblacional no difiera de una desviación estándar muestral por más de 2%?

6.48. La desviación estándar del tiempo de vida de 10 focos eléctricos que produce una compañía es igual a 120 horas. Encuentre los límites de confianza de *a*) 95% y *b*) 99% correspondientes a la desviación estándar de todos los focos que fabrica la compañía.

6.49. Resuelva el problema 6.48 si 25 focos muestran la misma desviación estándar de 120 horas.

6.50. Resuelva el problema 6.48 mediante el uso de la distribución χ^2 si una muestra de 100 focos denota la misma desviación estándar de 120 horas.

INTERVALOS DE CONFIANZA PARA RAZONES DE VARIANZAS

6.51. Las desviaciones estándar de los diámetros de cojinetes de bolas que producen dos máquinas son, respectivamente, de 0.042 cm y 0.035 cm, con base en muestras de tamaño 10 cada una. Determine los intervalos de confianza de *a*) 98% y *b*) 90% para la razón o cociente, de las varianzas.

6.52. Determine los intervalos de confianza de *a*) 98% y *b*) 90% para la razón de las desviaciones estándar del problema 6.51.

6.53. Dos muestras de tamaño 6 y 8 tienen la misma varianza. Calcule los intervalos de confianza de *a*) 98% y *b*) 90% para la razón de las varianzas de las poblaciones a las que corresponden.

6.54. Resuelva *a*) el problema 6.51 y *b*) el 6.53 si los tamaños de cada una de las muestras es de 120 unidades.

ESTIMACIONES DE LA MÁXIMA VEROSIMILITUD

6.55. Suponga que se hacen n observaciones $X_1, \ldots, X_n$ de una distribución de Poisson con parámetro λ desconocido. Encuentre la estimación de la máxima verosimilitud de λ.

6.56. Una población tiene una función de densidad dada por $f(x) = 2\nu\sqrt{\nu/\pi}x^2e^{-\nu x^2}, -\infty < x < \infty$. Para n observaciones, $X_1, \ldots, X_n$, que se hicieron de esta población, calcule la estimación de máxima verosimilitud de ν.

6.57. Una población tiene una función de densidad dada por

$$f(x) = \begin{cases} (k+1)x^k & 0 \le x \le 1 \\ 0 & \text{si no es así} \end{cases}$$

Para n observaciones $X_1, \ldots, X_n$ hechas de esta población, determine la estimación de la máxima verosimilitud de k.

PROBLEMAS DIVERSOS

6.58. Los coeficientes de confianza de 99% (dos colas) para la distribución normal están dados por $\pm$ 2.58. ¿Cuáles son los coeficientes correspondientes para la distribución t si *a*) $\nu = 4$, *b*) $\nu = 12$, *c*) $\nu = 25$, *d*) $\nu = 30$ y *e*) $\nu = 40$?

6.59. Una compañía tiene 500 cables. Una prueba de 40 cables seleccionados de manera aleatoria indican una resistencia a la ruptura media de 2 400 libras y una desviación estándar de 150 libras. *a*) ¿Cuáles son los límites de confianza de 95 y 99% para estimar la media de la resistencia a la ruptura de los 460 cables restantes? *b*) ¿Con qué grado de confianza es posible afirmar que la resistencia a la ruptura media de los 460 cables restantes corresponde a 2 400 $\pm$ 35 libras?

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

6.47. *a*) por lo menos 4 802 *b*) por lo menos 8 321 *c*) por lo menos 11 250

6.48. *a*) 87.0 a 230.9 horas *b*) 78.1 a 288.5 horas

6.49. *a*) 95.6 a 170.4 horas *b*) 88.9 a 190.8 horas

6.50. *a*) 106.1 a 140.5 horas *b*) 102.1 a 148.1 horas

6.51. *a*) 0.269 a 7.70 *b*) 0.453 a 4.58 **6.52.** *a*) 0.519 a 2.78 *b*) 0.673 a 2.14

6.53. *a*) 0.140 a 11.025 *b*) 0.264 a 5.124

6.54. *a*) 0.941 a 2.20, 1.067 a 1.944 *b*) 0.654 a 1.53, 0.741 a 1.35

6.55. $\lambda = \left(\sum x_k\right)/n$ **6.56.** $\nu = \dfrac{3n}{2(x_1^2 + \cdots + x_n^2)}$

6.57. $k = -1 - \dfrac{n}{\ln(x_1 \cdots x_n)}$

6.58. *a*) ±4.60 *b*) ±3.06 *c*) ±2.79 *d*) ±2.75 *e*) ±2.70

6.59. *a*) 2 400 ± 45 libras, 2 400 ± 59 libras *b*) 87.6%

PROBLEMAS APORTADOS

1. Se encuentra que la concentración promedio de zinc de una muestra de 36 cereales es de 2.6 gramos por miligramo. Encuentre los intervalos de confianza de *a*) 95% y *b*) 99% para la concentración media de zinc en el cereal. Suponga que la desviación estándar de la población es 0.3.

2. Un fabricante de reproductores de discos compactos utiliza un conjunto de pruebas amplias para evaluar la función eléctrica de su producto. Todos los reproductores de discos compactos deben pasar todas las pruebas antes de venderse. Una muestra aleatoria de 500 reproductores tiene como resultado 15 que fallan en una o más pruebas. Encuentre un intervalo de confianza de *a*) 90% y *b*) 95% para la proporción de los reproductores de discos compactos de la población que no pasarían todas las pruebas.

3. En un estudio de 300 accidentes de automóvil en una ciudad específica, 60 tuvieron consecuencias fatales. Con base en esta muestra, construya un intervalo de *a*) 95% y *b*) 99% de confianza para aproximar la proporción de todos los accidentes automovilísticos que en esa ciudad tienen consecuencias fatales.

CAPÍTULO 7

DECISIONES ESTADÍSTICAS

En la práctica, es frecuente tener que tomar decisiones acerca de poblaciones con base en información de muestreo. Tales decisiones se denominan *decisiones estadísticas*. Por ejemplo, quizá se desee decidir con base en datos de un muestreo si un nuevo suero es realmente eficaz para curar una enfermedad, si un procedimiento es mejor que otro o si una moneda está cargada.

HIPÓTESIS ESTADÍSTICAS. HIPÓTESIS NULAS

Cuando se deben tomar decisiones, es útil hacer suposiciones, o conjeturas, acerca de las poblaciones relacionadas. Las suposiciones, que pueden ser verdaderas o no, se llaman *hipótesis estadísticas* y, en general, son afirmaciones acerca de las distribuciones de probabilidad de las poblaciones.

Por ejemplo, si se quiere decidir si cierta moneda está cargada, se formula la hipótesis de que una moneda no está cargada, esto es, $p = 0.5$, donde p es la probabilidad de que ocurra una cara. De manera similar, si se quiere decidir si un procedimiento es mejor que otro, se formula la hipótesis de que *no hay diferencia* entre los procedimientos (esto es, cualquiera de las diferencias observadas son simplemente debidas a fluctuaciones en el muestreo de la *misma* población). A menudo, tales hipótesis se conocen como *hipótesis nulas* o, de manera sencilla, *hipótesis*, y se denotan por H_0.

Cualquier hipótesis que difiera de una hipótesis nula dada se llama *hipótesis alternativa*. Por ejemplo, si la hipótesis nula es $p = 0.5$, posibles hipótesis alternativas son $p = 0.7$, $p \neq 0.5$ o bien $p > 0.5$. Una hipótesis alternativa de la hipótesis nula se denota por H_1.

PRUEBAS DE HIPÓTESIS Y SIGNIFICANCIA

Si se supone que una hipótesis en particular es verdadera, y se encuentra que los resultados observados en una muestra aleatoria difieren notablemente de aquellos esperados bajo la hipótesis basada sólo en la coincidencia, con base en la teoría del muestreo, se puede decir que las diferencias observadas son *significativas* e inclinarse a descartar la hipótesis (o, por lo menos, no aceptarla con base en la evidencia obtenida). Por ejemplo, si 20 lanzamientos de la moneda producen 16 caras, sería posible descartar la hipótesis de que la moneda no está cargada, aunque podría ser una equivocación.

Los procedimientos que permiten decidir si se acepta o descarta una hipótesis o que determinan si las muestras observadas difieren de manera significativa de los resultados esperados se conocen como *pruebas de hipótesis, pruebas de significancia* o *reglas de decisiones.*

ERRORES DEL TIPO I Y DEL TIPO II

Si se rechaza una hipótesis cuando es cierta, se dice que se ha cometido un *error del tipo I*. Si, por otro lado, se acepta una hipótesis cuando debe descartarse, se dice que se ha cometido un *error del tipo II*. En cualquiera de los casos, se produce una decisión equivocada o un error de juicio.

Para que cualquier prueba de hipótesis o regla de decisión sea buena, debe diseñarse para minimizar errores de decisión. No es una cuestión sencilla, ya que en una muestra de cierto tamaño, como un intento para disminuir un tipo de error, se acompaña, en general, de un aumento en el otro tipo de error. En la práctica, un tipo de error puede

NIVEL DE SIGNIFICANCIA

Al comprobar una hipótesis, la probabilidad máxima que se estaría dispuesto a cometer un error del tipo 1 se llama *nivel de significancia* de la prueba. A menudo esta probabilidad se especifica antes de tomar una muestra, para que los resultados que se obtengan no influyan en la decisión.

En la práctica, se acostumbra que el nivel de significancia sea de 0.05 o 0.01, aunque se usan otros valores. Por ejemplo, si se elige un nivel de significancia de 0.05 o 5%, al diseñar la prueba de una hipótesis hay aproximadamente 5 posibilidades en 100 de que se descarte la hipótesis cuando debe aceptarse, esto es, siempre que la hipótesis sea verdadera tenemos una *confianza* de 95% de que se tomará la decisión correcta. En tales casos, se dice que la hipótesis ha sido *rechazada a un nivel de significancia de* 0.05, lo que quiere decir que es posible estar equivocado con una probabilidad de 0.05.

PRUEBAS EN LAS QUE INTERVIENE LA DISTRIBUCIÓN NORMAL

Para ilustrar las ideas anteriores, suponga que, con base en una hipótesis dada, la distribución de muestreo de un estadístico S es una distribución normal con media μ_S y una desviación estándar σ_S. Suponga también que se decide rechazar la hipótesis si S es demasiado pequeño o demasiado grande. La distribución de la variable estandarizada $Z = (S - \mu_S)/\sigma_S$ es la distribución normal estándar (media 0, varianza 1) que se muestra en la figura 7-1, por lo que los valores extremos de Z llevarían a rechazar la hipótesis.

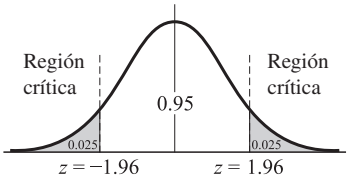


Figura 7-1

Como se indica en la figura, es posible tener una confianza de 95% de que, si la hipótesis fuera verdadera, el puntaje z de un estadístico real muestral S se encontraría entre -1.96 y 1.96 (puesto que el área debajo de la curva normal entre estos valores es 0.95).

Sin embargo, si al elegir una muestra de manera aleatoria se ve que el puntaje z de su estadístico se ubica *fuera* del rango -1.96 a 1.96, se debe concluir que tal evento podría suceder con la probabilidad de sólo 0.05 (el área sombreada total de la figura) si la hipótesis dada fuese verdadera. Se podría decir que este puntaje z difiere *de manera significativa* de lo que debería esperarse bajo la hipótesis, y se tendería a rechazar la hipótesis.

El área sombreada total, 0.05, es el nivel de significancia de la prueba, que representa la probabilidad de estar equivocado si se descarta una hipótesis, esto es, la probabilidad de cometer un error del tipo 1. Por tanto, se dice que la hipótesis fue *rechazada a un nivel de significancia de* 0.05, o que el puntaje z del estadístico muestral dado es *significativo* a un *nivel de significancia* de 0.05.

El conjunto del puntaje z fuera del rango de -1.96 a 1.96 constituye lo que se conoce como *región crítica*, *región de rechazo de la hipótesis* o *región de significancia*. En consecuencia, es posible denominar el conjunto del puntaje z dentro del rango de -1.96 a 1.96 como *región de aceptación de la hipótesis* o *región de no significancia.*

De acuerdo con los comentarios anteriores, se formula la siguiente regla de decisión:

a) Se debe rechazar la hipótesis a un nivel de significancia de 0.05 si el puntaje z del estadístico S está fuera del rango de -1.96 a 1.96 (esto es, o $z > 1.96$ o bien $z < -1.96$). Esto equivale a decir que el estadístico muestral observado es significativo a un nivel de 0.05.

b) En caso contrario, se debe aceptar la hipótesis (o, si se desea, no tomar ninguna decisión).

Se debe observar que pudieron usarse otros niveles de significancia. Por ejemplo, si se emplea un nivel de 0.01, se debería sustituir 1.96 por 2.58 en todos los lugares anteriores (vea la tabla 7-1). Se puede usar también la tabla 6-1, página 196, puesto que la suma del nivel de significancia y el nivel de confianza es 100%.

PRUEBAS DE UNA COLA Y DE DOS COLAS

En la prueba anterior se mostró interés en los valores extremos del estadístico S o en su puntaje z correspondiente en ambos lados de la media, esto es, en ambas colas de la distribución. Por tanto, estas pruebas se llaman *pruebas de dos colas* o *pruebas bilaterales*.

No obstante, muchas veces tal vez interese conocer sólo los valores del extremo de un lado de la media, esto es, en una cola de la distribución, como, por ejemplo, cuando se prueba la hipótesis de que un proceso es mejor que cualquier otro (lo que es diferente a probar si un proceso es mejor o peor que el otro). Estas pruebas se llaman *pruebas de una cola* o *pruebas unilaterales*. En estos casos, la región crítica es la región a un lado de la distribución, con área igual al nivel de significancia.

La tabla 7-1, que proporciona valores críticos de z de pruebas de una cola y de dos colas a varios niveles de significancia, será útil con fines de referencia. Los valores críticos de z en otros niveles de significancia se encuentran por medio de la tabla de áreas de curvas normales.

Tabla 7-1

Nivel de significancia α	0.10	0.05	0.01	0.005	0.002
Valores críticos de z para pruebas de una cola	-1.28 *o* 1.28	-1.645 *o* 1.645	–2.33 *o* 2.33	-2.58 *o* 2.58	-2.88 *o* 2.88
Valores críticos de z para pruebas de dos colas	-1.645 *y* 1.645	-1.96 *y* 1.96	–2.58 *y* 2.58	-2.81 *y* 2.81	-3.08 *y* 3.08

VALOR P

En la mayoría de las pruebas que se consideren, la hipótesis nula H_0 será una afirmación de que un parámetro poblacional tiene un valor específico, y la hipótesis alternativa H_1 corresponderá a una de las siguientes afirmaciones:

i) El parámetro es mayor que el valor indicado (prueba de la cola derecha).

ii) El parámetro es menor que el valor indicado (prueba de la cola izquierda).

iii) El parámetro es mayor o menor que el valor indicado (prueba de dos colas).

En los casos *i*) y *ii*), H_1 tiene una sola dirección con respecto al parámetro, y en el caso *iii*), H_1 es bidireccional. Después de emplear la prueba y calcular el estadístico S de la prueba, el valor P de ella es la probabilidad de que un valor de S sea en la dirección o direcciones de H_1 y tan extremo como el valor que ocurrió en la realidad debería ocurrir si H_0 fuera verdadera.

Por ejemplo, suponga que se sabe que la desviación estándar σ de una población normal es 3, y H_0 afirma que la media μ es igual a 12. Una muestra aleatoria de tamaño 36 que se toma de una población resulta en una media muestral de $\bar{x} = 12.95$. Se selecciona el siguiente estadístico de prueba:

$$Z = \frac{\bar{X} - 12}{\sigma/\sqrt{n}} = \frac{\bar{X} - 12}{0.5},$$

el cual, si H_0 es verdadera, es la variable aleatoria normal estándar. El valor de prueba de Z es $(12.95 - 12)/0.5 = 1.9$. En consecuencia, el valor P de la prueba depende de la hipótesis alternativa H_1 de acuerdo con lo siguiente:

i) Con H_1: $\mu > 12$ [caso *i*) anterior], el valor P es la probabilidad de que una muestra aleatoria de tamaño 36 resulte en una media muestral de 12.95, o mayor, si la media real fuera 12, esto es, $P(Z \geq 1.9) = 0.029$. En otras palabras, las posibilidades de que $\bar{x} > 12.95$ si $\mu = 12$ son alrededor de 3 en 100.

ii) Con H_1: $\mu < 12$ [caso *ii*) anterior], el valor P de la prueba es la probabilidad de que una muestra aleatoria de tamaño 36 resultaría en una media muestral de 12.95 o menor si la media real fuera 12, esto es, $P(Z \leq 1.9) = 0.97$, o bien las posibilidades de que $\bar{x} \leq 12.95$ si $\mu = 12$ son alrededor de 97 en 100.

iii) Con H_1: $\mu \neq 12$ [caso *iii*) anterior], el valor P es la probabilidad de que una muestra aleatoria de tamaño 36 resulte en una media muestral de 0.95 o mayor a partir de 12, esto es, $\bar{x} \geq 12.95$o bien $\bar{x} \leq 11.05$, si la media real fuera 12. Aquí el valor de P es $P(Z \geq 1.9) + P(Z \leq -1.9) = 0.057$, lo que indica que son aproximadamente 6 en 100 las probabilidades de que $|\bar{x} - 12| \geq 0.095$ si $\mu = 12$.

PRUEBAS ESPECIALES DE SIGNIFICANCIA EN EL CASO DE MUESTRAS GRANDES

En el caso de muestras grandes, muchos estadísticos S tienen casi las mismas distribuciones normales con media μ_S y desviación estándar σ_S. En estos casos, es posible usar los resultados anteriores para formular reglas de decisión o pruebas de hipótesis y significancia. Los siguientes casos especiales son unos cuantos ejemplos de los estadísticos de interés práctico. En cada caso los resultados se mantienen para poblaciones infinitas o para muestreo con reemplazo. Para muestreo sin reemplazo de poblaciones finitas, se deben modificar los resultados. Vea las páginas 156 y 158.

1. MEDIAS. En este caso, $S = \overline{X}$, la media muestral: $\mu_S = \mu_{\overline{X}} = \mu$, la media poblacional: $\sigma_S = \sigma_{\overline{X}} = \sigma/\sqrt{n}$, donde σ es la desviación estándar poblacional y n es el tamaño de la muestra. La variable estándar está dada por

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \tag{1}$$

Cuando es necesario, se usa la desviación estándar muestral observada, s (o bien $\hat{s}$) para estimar σ.

Para probar la hipótesis nula H_0 de que la media poblacional es $\mu = a$, se debería emplear el estadístico (1). En consecuencia, si la hipótesis alternativa es $\mu \neq a$, mediante una prueba de dos colas, se debería aceptar H_0 (o, por lo menos, no se rechaza) al nivel 0.05 si para una muestra de tamaño específico n con media $\bar{x}$

$$-1.96 \leq \frac{\bar{x} - a}{\sigma/\sqrt{n}} \leq 1.96 \tag{2}$$

y se debería rechazar en el caso contrario. En el caso de otros niveles de significancia, se debería cambiar (2) de manera adecuada. Para probar H_0 contra la hipótesis alternativa de que la media poblacional es mayor que a, se debería utilizar una prueba de una cola y aceptar H_0 (o, por lo menos, no se rechaza) en el nivel de 0.05 si

$$\frac{\bar{x} - a}{\sigma/\sqrt{n}} < 1.645 \tag{3}$$

(vea la tabla 7-1) y en otro caso se rechaza. Para probar H_0 contra la hipótesis alternativa de que la media poblacional es menor que a, se aceptaría H_0 en el nivel de 0.05 si

$$\frac{\bar{x} - a}{\sigma/\sqrt{n}} > -1.645 \tag{4}$$

2. PROPORCIONES. En este caso, $S = P$, la proporción de "éxitos" en una muestra; $\mu_S = \mu_P = p$, donde p es la proporción de población de éxitos y n es el tamaño de muestra; $\sigma_S = \sigma_P = \sqrt{pq/n}$, donde $q = 1 - p$. La variable estandarizada está dada por

$$Z = \frac{P - p}{\sqrt{pq/n}} \tag{5}$$

En el caso de $P = X/n$, donde X es el número real de éxitos en una muestra, (5) se convierte en

$$Z = \frac{X - np}{\sqrt{npq}} \tag{6}$$

Es posible hacer comentarios similares a los correspondientes a las pruebas de una y de dos colas en el caso de medias.

3. **DIFERENCIAS DE MEDIAS.** Sean $\overline{X}_1$ y $\overline{X}_2$ las medias muestrales que se obtuvieron de muestras grandes de tamaños n_1 y n_2 que se tomaron de poblaciones respectivas con medias μ_1 y μ_2 y desviaciones estándar σ_1 y σ_2. Considere la hipótesis nula de que *no hay diferencia* entre las medias poblacionales, esto es, $\mu_1 = \mu_2$. A partir de (11), página 157, al establecer $\mu_1 = \mu_2$, se observa que la distribución muestral de las diferencias de medias es aproximadamente normal con media y desviación estándar dadas por

$$\mu_{\bar{X}_1 - \bar{X}_2} = 0 \qquad \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{7}$$

donde, si es necesario, se pueden usar las desviaciones estándar muestrales observadas s_1 y s_2 (o bien $\hat{s}_1$ y $\hat{s}_2$) como estimaciones de σ_1 y σ_2.

Mediante el empleo de la variable estandarizada dada por

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - 0}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} \tag{8}$$

de manera similar a la descrita en la parte 1 anterior, es posible probar la hipótesis nula contra las hipótesis alternativas (o la significancia de una diferencia observada) a un nivel de significancia adecuado.

4. **DIFERENCIAS DE PROPORCIONES.** Sean P_1 y P_2 las proporciones muestrales que se obtuvieron en muestras grandes de tamaños n_1 y n_2 que se tomaron de poblaciones respectivas con proporciones p_1 y p_2. Considere la hipótesis nula de que *no hay diferencia* entre las proporciones de poblaciones, esto es, $p_1 = p_2$ y, en consecuencia, las muestras en realidad se toman de la misma población.

De (13), página 157, al dejar $p_1 = p_2 = p$, se observa que la distribución muestral de diferencias entre las proporciones es casi normal con media y desviación estándar dadas por

$$\mu_{P_1 - P_2} = 0 \qquad \sigma_{P_1 - P_2} = \sqrt{p(1 - p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \tag{9}$$

donde $\bar{P} = \dfrac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$ se usa como una estimación de la proporción poblacional p.

Al emplear la variable estandarizada

$$Z = \frac{P_1 - P_2 - 0}{\sigma_{P_1 - P_2}} = \frac{P_1 - P_2}{\sigma_{P_1 - P_2}} \tag{10}$$

se verifican las diferencias observadas a un nivel de significancia apropiado y, en consecuencia, la hipótesis nula.

Es factible idear de modo similar pruebas que relacionen otros estadísticos (vea la tabla 5-1, página 160).

PRUEBAS ESPECIALES DE SIGNIFICANCIA DE MUESTRAS PEQUEÑAS

En el caso de que las muestras sean pequeñas ($n < 30$), es posible formular pruebas de hipótesis y significancia mediante el empleo de otras distribuciones además de la normal, como la distribución t de Student, la ji-cuadrada y F, las cuales implican la teoría de muestreo exacta y, por tanto, es obvio que se mantienen iguales cuando las muestras son grandes, en cuyo caso se reducen a las dadas antes. A continuación se presentan algunos ejemplos.

1. **MEDIAS.** Para probar la hipótesis H_0 de que una población normal tiene media, μ, se emplea

$$T = \frac{\bar{X} - \mu}{S}\sqrt{n - 1} = \frac{\bar{X} - \mu}{\hat{S}}\sqrt{n} \tag{11}$$

$$= \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

con n grande, excepto que se usa $\hat{S} = \sqrt{n/(n-1)}\,S$ en lugar de σ. La diferencia es que mientras que Z está distribuida de manera normal, T tiene una distribución t de Student. A medida de que crece n, Z y T tienden a concordar. Es posible hacer pruebas de hipótesis similares a las de las medias de la página 216 mediante el uso de los valores críticos de t en lugar de los valores críticos de z.

2. **DIFERENCIAS DE MEDIAS.** Suponga que se toman dos muestras aleatorias de tamaños n_1 y n_2 de poblaciones normales (o casi normales), cuyas desviaciones estándar son iguales, esto es, $\sigma_1 = \sigma_2$. Suponga, además, que estas dos muestras tienen medias y desviaciones estándar dadas por $\bar{X}_1$, $\bar{X}_2$ y S_1, S_2, respectivamente. Para probar la hipótesis H_0 de que las muestras son de la misma población (esto es, tanto $\mu_1 = \mu_2$ como $\sigma_1 = \sigma_2$), se emplea la variable

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad \text{donde} \qquad \sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}} \tag{12}$$

La distribución de T es la distribución t de Student con $\nu = n_1 + n_2 - 2$ grados de libertad. Resulta factible el uso de (12) al hacer $\sigma_1 = \sigma_2 = \sigma$ en (12), página 157, y después usar como un estimador de σ_2 el promedio ponderado

$$\frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$$

donde $\hat{S}_1^2$ y $\hat{S}_2^2$ son los estimadores insesgados de σ_1^2 y σ_2^2. Ésta es la *varianza combinada* que se obtiene al combinar los datos.

3. **VARIANZAS.** Para probar la hipótesis H_0 de que una población normal tiene varianza σ^2, se consideran las variables aleatorias

$$\chi^2 = \frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \tag{13}$$

las cuales (vea las páginas 158-159) tienen la distribución ji-cuadrada con $n - 1$ grados de libertad. En consecuencia, si resulta que una muestra aleatoria de tamaño n tiene una varianza s^2, se debería aceptar H_0 con base en la prueba de dos colas (o al menos no se rechazaría) en el nivel de 0.05 si

$$\chi^2_{0.025} \leq \frac{ns^2}{\sigma^2} \leq \chi^2_{0.975} \tag{14}$$

y en otro caso se debería rechazar. Se obtiene un resultado similar con el nivel de 0.01 u otro.

Para probar la hipótesis H_1 de que la varianza poblacional es mayor que σ^2, se sigue usando la hipótesis nula H_0 pero ahora se emplea la prueba de una cola. Por tanto, se debería rechazar H_0 a un nivel de 0.05 (y así se concluye que H_1 es correcta) si la varianza muestral particular s^2 fuera tal que

$$\frac{ns^2}{\sigma^2} > \chi^2_{0.95} \tag{15}$$

y se debería aceptar H_0 (o al menos no se rechazaría) en otro caso.

4. **RAZONES DE VARIANZAS.** En algunos problemas se debe decidir si dos muestras de tamaños m y n cuyas varianzas medidas son s_1^2 y s_2^2, respectivamente, provienen o no de poblaciones normales con la misma varianza. En estos casos, se usa el estadístico (vea la página 159)

$$F = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2} \tag{16}$$

donde σ_1^2, σ_2^2 son las varianzas de las dos poblaciones normales de las que se toman las muestras. Suponga que H_0 denota la hipótesis nula de que no hay diferencia entre las varianzas poblacionales, esto es, $\sigma_1^2 = \sigma_2^2$. Entonces, con base en esta hipótesis (16) se convierte en

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} \tag{17}$$

Para probar esta hipótesis en el nivel de 0.10, por ejemplo, se nota primero que F en (16) tiene la distribución F con $m - 1$, $n - 1$ grados de libertad. En ese caso, con base en la prueba de dos colas, se debería aceptar H_0 (o no se rechazaría) en el nivel 0.10 si

$$F_{0.05} \leq \frac{\hat{s}_1^2}{\hat{s}_2^2} \leq F_{0.95} \tag{18}$$

y se debería rechazar en otro caso.

Se pueden formular enfoques similares si se aplican pruebas de una cola en caso de que se desee verificar la hipótesis de que cierta varianza poblacional es, en realidad, mayor que la otra.

RELACIÓN ENTRE LA TEORÍA DE LA ESTIMACIÓN Y LA PRUEBA DE HIPÓTESIS

A partir de los comentarios anteriores, resulta inevitable observar que hay una relación entre la teoría de la estimación que implica intervalos de confianza y la teoría de prueba de hipótesis. Por ejemplo, se observa que el resultado (2) para aceptar H_0 al nivel de 0.05 equivale al resultado (1) de la página 196, lo que conduce al intervalo de confianza de 95%

$$\bar{x} - \frac{1.96\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{1.96\sigma}{\sqrt{n}} \tag{19}$$

En consecuencia, por lo menos en el caso de las pruebas de dos colas, se pueden emplear los intervalos de confianza del capítulo 6 para probar hipótesis. Un resultado similar en el caso de las pruebas de una cola requeriría intervalos de confianza unilaterales (vea el problema 6.14).

CURVAS CARACTERÍSTICAS DE OPERACIÓN. POTENCIA DE UNA PRUEBA

Se ha explicado la forma de limitar el error del tipo I al elegir de manera apropiada el nivel de significancia. Es posible evitar de manera total el riesgo de los errores de tipo II simplemente al no hacerlos, lo que significa no aceptar nunca la hipótesis. No obstante, en muchos casos prácticos no es posible aplicar este enfoque. En tales casos, a menudo se usan las *curvas características de operación* o *curvas CO*, que son gráficas que muestran las probabilidades de errores de tipo II bajo varias hipótesis. Esto es, muestran qué tanto ayudan las pruebas a minimizar los errores, indican la *potencia de una prueba* para evitar tomar decisiones equivocadas. Son útiles para diseñar experimentos que muestren, por ejemplo, qué tamaño de muestra se debe usar.

GRÁFICAS DE CONTROL DE CALIDAD

Con frecuencia, en la práctica es de suma importancia saber si un proceso ha cambiado de manera suficiente como para que sea necesario tomar mejores medidas. Surgen, por ejemplo, problemas en el *control de calidad*, donde se tiene que decidir, muchas veces con premura, si los cambios observados se deben sólo a fluctuaciones aleatorias o a cambios reales en el proceso de fabricación por un deterioro de piezas de la maquinaria, o debido a errores de los empleados. Las *gráficas de control* brindan un método útil y sencillo para tratar estos problemas (vea el problema 7.29).

AJUSTE DE DISTRIBUCIONES TEÓRICAS A DISTRIBUCIONES DE FRECUENCIA MUESTRALES

Con frecuencia, cuando se tiene cierta indicación de la distribución de una población mediante razonamiento probabilístico o de otro tipo, es posible ajustar tales distribuciones teóricas (que también se conocen como distribuciones "modeladas" o "esperadas") respecto de distribuciones de frecuencia que se obtuvieron de una muestra de la población. Por lo general, el método que se utiliza consiste en emplear la media y la desviación estándar de la muestra para estimar las de la población. Vea los problemas 7.30, 7.32 y 7.33.

El problema de probar la *bondad de ajuste* de distribuciones teóricas a distribuciones muestrales es prácticamente igual al de decidir si hay diferencias de significancia entre los valores de la población y los de la muestra. Se describe a continuación la *prueba ji-cuadrada*, una prueba de significancia importante para la bondad de ajuste de distribuciones teóricas.

Cuando se intenta determinar si una distribución normal representa un ajuste preciso de los datos dados, es conveniente usar *papel de graficación de la curva normal*, o *papel para graficación de probabilidad*, como a veces se le llama (vea el problema 7.31).

LA PRUEBA JI-CUADRADA PARA LA BONDAD DE AJUSTE

Para determinar si una proporción P de "éxitos" en una muestra de tamaño n que se toma de una población binomial difiere de manera significativa de la proporción poblacional p de éxitos, se utilizó el estadístico dado por (5) o (6) en las páginas 216 y 217. En este caso sencillo, es posible que sucedan sólo dos eventos, A_1 y A_2, que se han llamado "éxito" y "fracaso" y que tienen las probabilidades p y $q = 1 - p$, respectivamente. A menudo, un valor de muestreo particular de la variable aleatoria $X = nP$ se llama *frecuencia observada* del evento A_1, mientras que np se conoce como la *frecuencia esperada.*

EJEMPLO 7.1 Si se obtiene una muestra de 100 lanzamientos de una moneda no cargada, se tiene que $n = 100$ $p = \frac{1}{2}$, entonces la frecuencia esperada de caras (éxitos) es $np = (100)(\frac{1}{2}) = 50$. Por supuesto que la frecuencia observada en la muestra podría ser diferente.

Una generalización natural se aplica al caso donde pueden ocurrir k eventos posibles, $A_1, A_2, \ldots, A_k$, con las probabilidades respectivas $p_1, p_2, \ldots, p_k$. En estos casos se tiene una *población multinomial* (vea la página 112). Si se toma una muestra de tamaño n de esta población, es posible describir las frecuencias observadas de los eventos $A_1, A_2, \ldots, A_k$ con variables aleatorias $X_1, \ldots, X_k$ (cuyos valores específicos $x_1, x_2, \ldots, x_k$ serían las frecuencias observadas de la muestra), mientras que las frecuencias esperadas estarían dadas por $np_1, \ldots, np_k$, respectivamente. Los resultados pueden indicarse como se muestra en la tabla 7-2.

Tabla 7-2

Evento	A_1	A_2	$\cdots$	A_k
Frequencia observada	x_1	x_2	$\cdots$	x_k
Frecuencia esperada	np_1	np_2	$\cdots$	np_k

EJEMPLO 7.2 Si se obtiene una muestra de 120 lanzamientos de un dado no cargado tal que $n = 120$, las probabilidades de las caras 1, 2, . . . , 6 se denotan mediante $p_1, p_2, \ldots, p_6$, respectivamente, y todas son iguales a $\frac{1}{6}$. Las frecuencias esperadas son $np_1, np_2, \ldots, np_6$ y todas son iguales a $(120)(\frac{1}{6}) = 20$. Es claro que las frecuencias observadas de las diversas caras que aparecen en la muestra pueden ser diferentes.

Una clave de la posible generalización del estadístico (6), que podría medir las discrepancias existentes entre las frecuencias observadas y las esperadas en la tabla 7-2, se obtiene al elevar al cuadrado el estadístico (6) y escribirlo como

$$Z^2 = \frac{(X - np)^2}{npq} = \frac{(X_1 - np)^2}{np} + \frac{(X_2 - nq)^2}{nq} \tag{20}$$

donde $X_1 = X$ es la variable aleatoria asociada con "éxito" y $X_2 = n - X_1$ es la variable aleatoria asociada con "fracaso". Observe que nq en (20) es la frecuencia esperada de fracasos.

La forma del resultado (20) sugiere que una medida de la discrepancia que se observa entre las frecuencias observadas y las esperadas en el caso general se proporciona mediante el estadístico

$$\chi^2 = \frac{(X_1 - np_1)^2}{np_1} + \frac{(X_2 - np_2)^2}{np_2} + \cdots + \frac{(X_k - np_k)^2}{np_k} = \sum_{j=1}^{k} \frac{(X_j - np_j)^2}{np_j} \tag{21}$$

donde la frecuencia total (esto es, el tamaño de la muestra) es n, por lo que

$$X_1 + X_2 + \cdots + X_k = n \tag{22}$$

Una expresión equivalente a (21) es

$$\chi^2 = \sum_{j=1}^{k} \frac{X_j^2}{np_j} - n \tag{23}$$

Si $\chi^2 = 0$, las frecuencias observadas y las esperadas concuerdan con exactitud, mientras que la concordancia no ocurre si $\chi^2 > 0$, Cuanto más grande es el valor de χ^2, tanto mayor es la discrepancia entre las frecuencias observadas y las esperadas.

Como se muestra en el problema 7.62, la distribución muestral de χ^2, como se define en (21), se aproxima muy estrechamente a la distribución ji-cuadrada [de ahí la elección del símbolo en (21)] si las frecuencias esperadas np_j equivalen por lo menos a 5; la aproximación mejora en el caso de valores más grandes. El número de grados de libertad de esta distribución ji-cuadrada está dado por:

a) $\nu = k - 1$ si las frecuencias esperadas se pueden calcular sin tener que estimar los parámetros poblacionales de los estadísticos muestrales, o de muestra. Observe que si resta 1 de k debido a la condición de restricción (22), que señala que si se conoce $k - 1$ de las frecuencias esperadas es posible determinar la frecuencia restante.

b) $\nu = k - 1 - m$ si es posible calcular las frecuencias esperadas estimando sólo m parámetros poblacionales de los estadísticos muestrales, o de muestreo.

En la práctica, se calculan frecuencias esperadas con base en la hipótesis H_0. Si con base en esta hipótesis, el valor calculado de χ^2 dado por (21) o (23) es más grande que algún valor crítico (como $\chi^2_{0.95}$ o bien $\chi^2_{0.99}$, que son los valores críticos a los niveles de significancia de 0.05 y 0.01, respectivamente), se debería concluir que las frecuencias observadas difieren de manera *significativa* de las frecuencias esperadas y se rechazaría H_0 a un nivel de significancia correspondiente. Si el caso es el contrario, se aceptaría o, por lo menos, no se rechazaría. Este procedimiento se llama *prueba de ji-cuadrada* de la hipótesis o significancia.

Además de aplicar la distribución multinomial, es posible usar la prueba ji-cuadrada para determinar qué tan bien se ajustan otras distribuciones teóricas, como la distribución normal o la de Poisson, a distribuciones empíricas, es decir, aquellas que se obtienen de los datos de muestreos. Vea el problema 7.44.

TABLAS DE CONTINGENCIA

La tabla 7-2 anterior, donde las frecuencias observadas ocupan una sola fila, se denomina *tabla de clasificación unidireccional*. Puesto que el número de columnas es k, también se llama *tabla* de $1 \times k$ (que se lee "1 por k"). Al extender estas ideas, es posible llegar a las *tablas de clasificación bidireccional*, o *tablas* $h \times k$, en las cuales las frecuencias observadas ocupan h filas y k columnas. Tales tablas también se conocen como *tablas de contingencia.*

En correspondencia a cada frecuencia observada en una tabla de contingencia $h \times k$ hay una *frecuencia esperada* o *teórica*, la cual se calcula con base en alguna hipótesis de acuerdo con las reglas de la probabilidad. Estas frecuencias que ocupan las *celdas* de una tabla de contingencia se llaman *frecuencias de las celdas*. La frecuencia total en cada fila o cada columna se llama *frecuencia marginal.*

Para investigar la concordancia entre las frecuencias observadas y las esperadas, se calculó el estadístico

$$\chi^2 = \sum_j \frac{(X_j - np_j)^2}{np_j} \tag{24}$$

donde se calcula la suma de todas las celdas de la tabla de contingencia, y los símbolos X_j y np_j representan, respectivamente, las frecuencias observadas y las esperadas en la celda j-ésima. Esta suma, que es análoga a (21), contiene hk términos. La suma de todas las frecuencias observadas se denota n y es igual a la suma de todas las frecuencias esperadas [compare con la ecuación (22)].

Como antes, el estadístico (24) tiene una distribución muestral muy próxima a la distribución ji-cuadrada siempre que las frecuencias esperadas no sean demasiado pequeñas. El número de grados de libertad ν de esta distribución ji-cuadrada para $h > 1$, $k < 1$ se determina así

a) $\nu = (h - 1)(k - 1)$ si las frecuencias esperadas pueden calcularse sin tener que estimar parámetros poblacionales de los estadísticos muestrales, o del muestreo. Para realizar una prueba de lo anterior, vea el problema 7.48.

b) $\nu = (h - 1)(k - 1) - m$ si es posible calcular las frecuencias esperadas sólo mediante la estimación de los m parámetros poblacionales de los estadísticos muestrales, o del muestreo.

Las pruebas de significancia de las tablas $h \times k$ son parecidas a las de las tablas $1 \times k$. Las frecuencias esperadas se sujetan a una hipótesis H_0 particular. Una hipótesis que se prueba a menudo es la de que dos clasificaciones son independientes una de la otra.

Es posible extender las tablas de contingencia a dimensiones superiores. Por ejemplo, se pueden tener tablas $h \times k \times l$ donde están presentes tres clasificaciones.

CORRECCIÓN DE YATES PARA LA CONTINUIDAD

Cuando se aplican los resultados de distribuciones continuas a datos discretos, se pueden hacer ciertas correcciones de continuidad, como se ha observado en capítulos anteriores. Se cuenta con una corrección similar cuando se usa la distribución ji-cuadrada. La corrección consiste en reescribir (21) como

$$\chi = \frac{(|X_1 - np_1| - 0.5)^2}{np_1} + \frac{(|X_2 - np_2| - 0.5)^2}{np_2} + \cdots + \frac{(|X_k - np_k| - 0.5)^2}{np_k} \tag{25}$$

y a menudo se refiere como la *corrección de Yates*. También existe una modificación análoga de (24).

En general, se hace la corrección sólo cuando el número de grados de libertad es $\nu = 1$. En el caso de muestras grandes, esto produce casi los mismos resultados que los de la χ^2 no corregida, pero pueden surgir dificultades cerca de los valores críticos (vea el problema 7.41). En el caso de muestras pequeñas donde cada frecuencia esperada está entre 5 y 10, quizá sea mejor comparar los valores corregidos y no corregidos de χ^2. Si ambos valores conducen a la misma conclusión en cuanto a la hipótesis, como el rechazo a un nivel de 0.05, es raro que se encuentren dificultades. Si dan lugar a conclusiones diferentes, se pueden aumentar los tamaños de las muestras o, si este paso no es práctico, es posible emplear métodos de probabilidad exactos que implican a la distribución multinomial.

COEFICIENTE DE CONTINGENCIA

Una medida del grado de la relación, asociación o dependencia de las clasificaciones en una tabla de contingencia está dada por

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} \tag{26}$$

lo que se llama *coeficiente de contingencia*. Cuanto más grande es el valor de C, tanto mayor es el grado de asociación. El número de filas y columnas de la tabla de contingencia determina el valor máximo de C, el cual nunca es mayor que uno. En el caso de una tabla $k \times k$, el valor máximo de C está dado por $\sqrt{(k-1)/k}$. Vea los problemas 7.52 y 7.53.

PROBLEMAS RESUELTOS

PRUEBAS DE MEDIAS Y PROPORCIONES UTILIZANDO DISTRIBUCIÓN NORMAL

7.1. Determinar la probabilidad de lograr entre 40 y 60 caras, inclusive, en 100 lanzamientos de una moneda no cargada.

De acuerdo con la distribución binomial, la probabilidad que buscamos es

$${}_{100}C_{40}\left(\frac{1}{2}\right)^{40}\left(\frac{1}{2}\right)^{60} + {}_{100}C_{41}\left(\frac{1}{2}\right)^{41}\left(\frac{1}{2}\right)^{59} + \cdots + {}_{100}C_{60}\left(\frac{1}{2}\right)^{60}\left(\frac{1}{2}\right)^{40}$$

La media y la desviación estándar del número de caras en 100 lanzamientos están dadas por

$$\mu = np = 100\left(\frac{1}{2}\right) = 50 \qquad \sigma = \sqrt{npq} = \sqrt{(100)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 5$$

Puesto que tanto np como nq son mayores que 5, se puede usar la aproximación normal a la distribución binomial para evaluar la suma anterior.

En una escala continua, entre 40 y 60 caras, inclusive, es equivalente a entre 39.5 y 60.5 caras.

$$39.5 \text{ en unidades estándar} = \frac{39.5 - 50}{5} = -2.10 \qquad 60.5 \text{ en unidades estándar} = \frac{60.5 - 50}{5} = 2.10$$

$$\text{Probabilidad que se busca} = \text{área bajo de la curva normal entre } z = -2.10 \text{ y } z = 2.10$$
$$= 2(\text{área entre } z = 0 \text{ y } z = 2.10) = 2(0.4821) = 0.9642$$

7.2. Para probar la hipótesis de que una moneda no está cargada, se adoptan las siguientes reglas de decisión: 1) aceptar la hipótesis si el número de caras en una sola muestra de 100 lanzamientos está entre 40 y 60, inclusive, 2) rechazar la hipótesis en otro caso.

a) Encontrar la probabilidad de rechazar la hipótesis cuando en la realidad es la hipótesis correcta.

b) Interpretar de manera gráfica la regla de decisión y el resultado del inciso *a*).

c) ¿A cuáles conclusiones se puede llegar si la muestra de 100 lanzamientos resultaran 53 caras? ¿60 caras?

d) ¿Se puede estar equivocado en las conclusiones de *c*)? Explicar la respuesta.

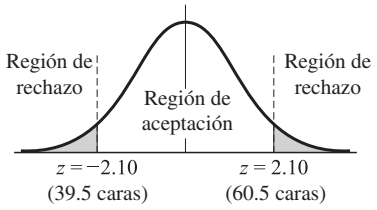


Figura 7-2

Si una sola muestra de 100 lanzamientos produce un puntaje z entre -2.10 y 2.10, se acepta la hipótesis; en caso contrario se rechaza y se decide que la moneda está cargada.

El error que se comete al rechazar la hipótesis cuando ésta se debe aceptar es el *error del tipo I* de la regla de decisión; y la probabilidad de cometer este error, que es igual a 0.0358 de la parte *a*), se representa mediante el área sombreada total de la figura.

Si una sola muestra de 100 lanzamientos produce un número de caras cuyo puntaje z está en las regiones sombreadas, se puede decir que este puntaje z difiere *de manera significativa* del esperado si la hipótesis fuera verdadera. Por tanto, el área total sombreada (esto es, la probabilidad de un error del tipo I) se conoce como *nivel de significancia* de la regla de decisión; en este caso equivale a 0.0358. Así que se habla de rechazar la hipótesis a un nivel de significancia de 0.0358 o 3.58%.

c) De acuerdo con la regla de decisión, se tendría que aceptar la hipótesis de que la moneda no está cargada en ambos casos. Se podría argumentar que si sólo se hubiera obtenido una cara más, se habría rechazado la hipótesis. Esto es lo que se tiene que enfrentar cuando se usa cualquier línea de división fina al tomar una decisión.

d) Sí. Podría aceptarse la hipótesis cuando ésta debe en realidad rechazarse, como sería el caso, por ejemplo, cuando la probabilidad de caras en realidad es de 0.7 en vez de 0.5.

El error cometido en aceptar la hipótesis cuando ésta se debe rechazar es el *error del tipo II* de la decisión. Para mayor análisis, vea los problemas 7.23 a 7.25.

7.3. Diseñar una regla de decisión para probar la hipótesis de que una moneda no está cargada si se toma una muestra de 64 lanzamientos y si se usa el nivel de significancia de *a*) 0.05, *b*) 0.01.

a) **Primer método**

Si el nivel de significancia es de 0.05, cada área sombreada en la figura 7-3 es de 0.025 por simetría. En consecuencia, el área entre 0 y z_1 es $0.5000 - 0.0250 = 0.4750$ y $z_1 = 1.96$.

Por tanto, una posible regla de decisión es:

1) Aceptar la hipótesis que la moneda no está cargada si Z está entre -1.96 y 1.96.
2) Rechazar la hipótesis en caso contrario.

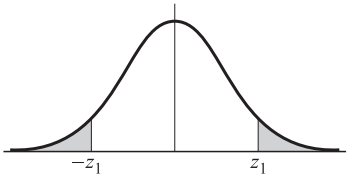


Figura 7-3

Los valores críticos -1.96 y 1.96 también se pueden encontrar en la tabla 7-1.

Para expresar esta regla de decisión en términos del número de caras que se obtiene en 64 lanzamientos de la moneda, observe que la media y desviación estándar de la distribución binomial exacta de caras están dadas por

$$\mu = np = 64(0.5) = 32 \qquad \text{y} \qquad \sigma = \sqrt{npq} = \sqrt{64(0.5)(0.5)} = 4$$

bajo la hipótesis de que la moneda no está cargada. Entonces $Z = (X - \mu)/\sigma = (X - 32)/4$.

$= 1.96$, $(X - 32)/4 = 1.96$ o sea $X = 39.84$. Si $Z = -1.96$, $(X - 32)/4 = -1.96$ o sea $X = 24.16$.

En consecuencia, la regla de decisión es:

1) Aceptar la hipótesis de que la moneda no está cargada si el número de caras se ubica entre 24.16 y 39.84, esto es, entre 25 y 39 inclusive.
2) Rechazar la hipótesis si no es así.

Segundo método

Con probabilidad de 0.95, el número de caras yacerá entre $\mu - 1.96\sigma$ y $\mu + 1.96\sigma$, esto es, $np - 1.96\sqrt{npq}$ y $np + 1.96\sqrt{npq}$ o entre $32 - 1.96(4) = 24.16$ y $32 + 1.96(4) = 39.84$, lo que conduce a la regla de decisión anterior.

Tercer método

a) $-1.96 < Z < 1.96$ equivale a $-1.96 < (X - 32)/4$. En consecuencia, $-1.96(4) < (X - 32) < 1.96(4)$ o bien $32 - 1.96(4) < X < 32 + 1.96(4)$, esto es, $24.16 < X < 39.84$, lo que también lleva a la regla de decisión anterior.

b) Si el nivel de significancia es 0.01, cada área sombreada de la figura 7-3 es 0.005. Entonces, el área entre 0 y z_1 es $0.5000 - 0.0050 = 0.4950$ y $z_1 = 2.58$ (más precisamente, 2.575). También es posible leer este resultado en la tabla 7-1.

De acuerdo con el procedimiento del segundo método de la parte *a*), se observa que con una probabilidad de 0.99 el número de caras estará entre $\mu - 2.58\sigma$ y $\mu + 2.58\sigma$, esto es, $32 - 2.58(4) = 21.68$ y $32 + 2.58(4) = 42.32$.

Por tanto, la regla de decisión es:

1) Aceptar la hipótesis si el número de caras está entre 22 y 42 inclusive.
2) En caso contrario, rechazar la hipótesis.

7.4. ¿Cómo se puede diseñar una regla de decisión en el problema 7.3 para evitar un error del tipo II?

Cometemos un error del tipo II si aceptamos la hipótesis cuando ésta debe rechazarse. Para evitar este error, en vez de aceptar la hipótesis, sencillamente no se la rechaza, lo que puede significar que, en este caso, se aplaza cualquier decisión. Por ejemplo, es posible enunciar de otra manera la regla de decisión para el problema 7.3*b*):

1) No rechazar la hipótesis si el número de caras está entre 22 y 42 inclusive.
2) Rechazar la hipótesis en otro caso.

No obstante, en muchos casos prácticos es importante decidir si debemos aceptar o rechazar una hipótesis. Un análisis completo de tales casos requiere la consideración de errores del tipo II (vea los problemas 7.23 a 7.25).

7.5. En un experimento sobre percepción extrasensorial (PES) se pide a un sujeto sentado en un cuarto indicar el color (rojo o azul) de una carta seleccionada de una baraja de 50 cartas bien barajadas por un individuo ubicado en otro cuarto. El primer sujeto no sabe cuántas cartas rojas y azules hay en la baraja. Si identifica correctamente 32 cartas, determinar si los resultados son significativos a un nivel de significancia de *a*) 0.05, *b*) 0.01. *c*) Encontrar e interpretar el valor *P* de la prueba.

Si p es la probabilidad de que el sujeto indique de manera correcta el color de una carta, entonces se debe decidir entre las dos hipótesis siguientes:

H_0: $p = 0.5$, y el sujeto nada más está adivinando, esto es, los resultados se deben a una coincidencia.

H_1: $p > 0.5$, y el sujeto tiene poderes extrasensoriales.

Elegimos una prueba de una cola, puesto que no es de interés la capacidad de obtener calificaciones muy bajas sino más bien la de lograr calificaciones altas.

Si la hipótesis H_0 es verdadera, la media y desviación estándar del número de cartas identificadas correctamente están dadas por

$$\mu = np = 50(0.5) = 25 \qquad \text{y} \qquad \sigma = \sqrt{npq} = \sqrt{50(0.5)(0.5)} = \sqrt{12.5} = 3.54$$

a) En el caso de una prueba de una cola a un nivel de significancia de 0.05, debemos elegir z_1 de la figura 7-4, de modo que el área sombreada en la región crítica de calificaciones altas sea 0.05. En consecuencia, el área entre 0 y z_1 es 0.4500, y $z_1 = 1.645$. También se puede leer este resultado en la tabla 7-1.

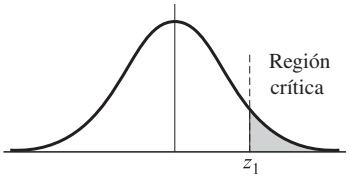


Figura 7-4

Por tanto, la regla de decisión o prueba de significancia es:

1) Si se observa que el puntaje z es mayor que 1.645, los resultados son significativos a un nivel de 0.05 y el individuo tiene poderes extrasensoriales.
2) Si el puntaje z es menor que 1.645, los resultados se deben a la casualidad, esto es, no son significativos a un nivel de 0.05.

Debido a que 32 en unidades estándar es $(32 - 25)/3.54 = 1.98$, que es mayor que 1.645, la decisión 1) se mantiene, esto es, se concluye que a un nivel de 0.05 el individuo tiene PES.

Observemos que se debe aplicar una corrección de continuidad, puesto que 32 en una escala continua está entre 31.5 y 32.5. Sin embargo, 31.5 tiene una calificación estándar de $(31.5 - 25)/3.54 = 1.84$, por lo que se llega a la misma conclusión.

b) Si el nivel de significancia es 0.01, entonces el área entre 0 y z_1 es 0.4900, y $z_1 = 2.33$. Puesto que 32 (o bien 31.5) en unidades estándar es 1.98 (o 1.84), que es menor que 2.33, se concluye que los resultados *no son significativos* al nivel de 0.01.

Algunos estadísticos adoptan la terminología de *altamente significativo* para designar resultados significativos a un nivel de 0.01, *probablemente significativo* para resultados significativos a un nivel de 0.05 pero no al de 0.01, y *no significativo* para resultados significativos a niveles mayores que 0.05.

De acuerdo con esta terminología, se tendría que concluir que los resultados experimentales anteriores son *probablemente significativos*, por lo que es posible que se justifiquen investigaciones adicionales del fenómeno.

c) El valor *P* de la prueba es la probabilidad de que se identificarían correctamente los colores de 32 o más cartas en una selección aleatoria. La calificación estándar de 32, tomando en cuenta la corrección de continuidad, es $z = 1.84$. En consecuencia, el valor *P* es $P(Z \geq 1.84) = 0.032$. El estadístico podría decir que con base en el experimento las probabilidades de estar equivocado al concluir que el individuo tiene poderes de PES son alrededor de 3 en 100.

7.6. El fabricante de una medicina de patente afirmó que la misma fue 90% eficaz para aliviar una alergia durante un periodo de 8 horas. En una muestra de 200 personas que padecían de la alergia, la medicina proporcionó alivio a 160 personas.

a) Determinar si la afirmación del fabricante es legítima con base en un nivel de significancia de 0.01.

b) Encontrar el valor *P* de la prueba.

a) Considere que p denota la probabilidad de experimentar alivio de la alergia al usar la medicina. En consecuencia, es necesario decidir entre las dos hipótesis:

$$H_0: p = 0.9, \text{ y la afirmación es correcta,}$$
$$H_1: p < 0.9, \text{ y la afirmación es falsa}$$

Elegimos una prueba de una cola, puesto que es de interés determinar si la proporción de personas aliviadas por la medicina es demasiado baja.

Si el nivel de significancia que se emplea es 0.01, esto es, si el área sombreada en la figura 7-5 es 0.01, entonces $z_1 = -2.33$, como vimos en el problema 7.5*b*) al usar la simetría de la curva, o de la tabla 7-1.

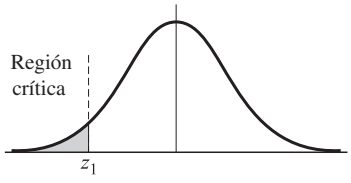


Figura 7-5

$$\mu = np = 200(0.9) = 180 \text{ y } \sigma = \sqrt{npq} = \sqrt{(200)(0.9)(0.1)} = 4.23.$$

Ahora, 160 en unidades estándar es $(160 - 180)/4.23 = -4.73$, que es mucho menor que -2.33. Por tanto, por la regla de decisión concluimos que la afirmación no es legítima y que los resultados muestrales son *altamente significativos* (vea el final del problema 7.5).

b) El valor P de la prueba es $P(Z \leq -4.73) \approx 0$. Esto muestra que es casi seguro que la afirmación es falsa. Si H_0 fuera verdadera, es casi seguro que la muestra aleatoria de 200 personas que padecen de alergias y usan la medicina incluiría más de 160 personas que experimentaron alivio.

7.7. Se calcula que la media del tiempo de vida de una muestra de 100 focos fluorescentes producidos por una compañía será de 1 570 horas con una desviación estándar de 120 horas. Si μ es la media del tiempo de vida de todos los focos que fabrica la compañía, pruebe la hipótesis $\mu = 1\,600$ horas contra la hipótesis alternativa $\mu \neq 1\,600$ horas, utilizando el nivel de significancia de *a*) 0.05 y *b*) 0.01. *c*) Determinar el valor P de la prueba.

Se tiene que decidir entre dos hipótesis:

$$H_0: \mu = 1\,600 \text{ horas} \qquad H_1: \mu \neq 1\,600 \text{ horas}$$

Usaremos una prueba de dos colas ya que $\mu \neq 1\,600$ incluye valores tanto mayores como menores que 1 600.

a) En el caso de una prueba de dos colas a un nivel de significancia de 0.05, se tiene la siguiente regla de decisión:

1) Rechazar H_0 si el puntaje z de la media muestral está fuera del rango -1.96 a 1.96.
2) Aceptar H_0 (o aplazar una decisión) si éste no es el caso.

El estadístico que se considera es la media muestral $\bar{X}$. La distribución muestral X tiene una media de $\mu_{\bar{X}} = \mu$ y una desviación estándar de $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, donde μ y σ son la media y desviación estándar de la población de todos los focos que produce la compañía.

Con base en la hipótesis H_0, se tiene $\mu = 1\,600$ a $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 120/\sqrt{100} = 12$, usando la desviación estándar muestral como una estimación de σ. Puesto que $Z = (\bar{X} - 1\,600)/12 = (1\,570 - 1\,600)/12 = -2.50$ está fuera del rango de -1.96 a 1.96, se rechaza H_0 a un nivel de significancia de 0.05.

b) Si el nivel de significancia es 0.01, se sustituye el rango de -1.96 a 1.96 en la regla de decisión de la parte *a*) por el de -2.58 a 2.58. Entonces, puesto que el puntaje z de -2.50 está dentro de este rango, se acepta H_0 (o se aplaza cualquiera decisión) a un nivel de significancia 0.01.

c) El valor P de la prueba de dos colas es $P(Z \leq -2.50) + P(Z \geq 2.50) = 0.0124$, el cual corresponde a la probabilidad de que la media del tiempo de vida sea menor que 1 570 horas o mayor que 1 630 horas ocurriría al azar si H_0 fuera verdadera.

7.8. En el problema 7.7 probar la hipótesis $\mu = 1\,600$ horas contra la hipótesis alternativa $\mu < 1\,600$ horas usando un nivel de significancia de *a*) 0.05, *b*) 0.01. *c*) Encontrar el valor P de la prueba.

Se debe decidir entre dos hipótesis

$$H_0: \mu = 1\,600 \text{ horas} \qquad H_1: \mu < 1\,600 \text{ horas}$$

Aquí se tiene que usar una prueba de una cola (vea la figura 7-5).

a) Si el nivel de significancia es 0.05, la región sombreada de la figura 7-5 tiene un área de 0.05, y se calcula que $z_1 = -1.645$. Por tanto, adoptamos la regla de decisión:

1) Rechazar H_0 si Z es menor que -1.645.
2) Aceptar H_0 (o aplazar la decisión) en otro caso.

Como en el problema 7.7*a*), puesto que el puntaje z es -2.50, lo que es menor que -1.645, se rechaza H_0 a un nivel de significancia de 0.05. Observamos que esta decisión es idéntica a la del problema 7.7*a*) cuando se usa la prueba de dos colas.

$$= \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{1\,850 - 1\,800}{100/\sqrt{50}} = 3.55$$

que es mayor que 2.33. En consecuencia, se concluye que los resultados son *altamente significativos* y se debe respaldar la afirmación.

b) El valor *P* de la prueba es $P(Z \geq 3.55) = 0.0002$, que es la probabilidad de que la media de la resistencia a la ruptura de 1 850 libras o más ocurriría por casualidad si H_0 fuera verdadera.

PRUEBAS EN LAS QUE INTERVIENEN LAS DIFERENCIAS DE MEDIAS Y PROPORCIONES

7.10. Se aplicó un examen a dos grupos de alumnos, el primero con 40 y el segundo con 50 estudiantes. En el primer grupo, la calificación media fue de 74 con una desviación estándar de 8, mientras que el segundo tenía una calificación media de 78 y una desviación estándar de 7. ¿Hay diferencia significativa entre el desempeño de las dos clases a un nivel de significancia de *a*) 0.05, *b*) 0.01? *c*) ¿Cuál es el valor *P* de la prueba?

Suponemos que los dos grupos son de dos poblaciones diferentes con medias respectivas μ_1 y μ_2. Tenemos que decidir entre las hipótesis

H_0: $\mu_1 = \mu_2$ y la diferencia se debe simplemente al azar
H_0: $\mu_1 \neq \mu_2$ y hay una diferencia significativa entre las dos clases

Bajo la hipótesis H_0, ambos grupos provienen de la misma población. La media y la desviación estándar de la diferencia de medias están dadas por

$$\mu_{\bar{X}_1 - \bar{X}_2} = 0 \qquad \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{8^2}{40} + \frac{7^2}{50}} = 1.606$$

donde se utilizaron las desviaciones estándar muestrales como estimaciones de σ_1 y σ_2.

$$= \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{74 - 78}{1.606} = -2.49$$

a) En el caso de una prueba de dos colas, los resultados son significativos a un nivel de 0.05 si Z yace fuera del rango de -1.96 a 1.96. Por tanto, se concluye que a un nivel de 0.05 hay una diferencia significativa en el desempeño de las dos clases y que es probable que la segunda clase sea mejor.

b) En el caso de una prueba de dos colas, los resultados son significativos a un nivel de 0.01 si Z se ubica fuera del rango de -2.58 a 2.58, por lo que se concluye que a un nivel de 0.01 no hay una diferencia significativa entre las clases.

Puesto que los resultados son significativos a un nivel de 0.05 pero no a uno de 0.01, se concluye que los resultados *probablemente son significativos* de acuerdo con la terminología que se emplea al final del problema 7.5.

c) El valor P de la prueba de dos colas es $P(Z \leq -2.49) + P(Z \geq 2.49) = 0.0128$, que es la probabilidad de que los estadísticos observados ocurrieran en la misma población.

7.11. La estatura media de 50 estudiantes hombres que mostraron mejores resultados que el promedio en los juegos atléticos escolares, fue de 68.2 pulgadas con una desviación estándar de 2.5 pulgadas, mientras que 50 estudiantes hombres que no mostraron interés en participar tenían una estatura media de 67.5 pulgadas con una desviación estándar de 2.8 pulgadas. *a*) Probar la hipótesis de que los estudiantes hombres que participan en deportes escolares son más altos que los demás estudiantes. *b*) ¿Cuál es el valor P de la prueba?

a) Se tiene que decidir entre dos hipótesis:

H_0: $\mu_1 = \mu_2$ y no hay diferencia entre las estaturas medias

H_1: $\mu_1 > \mu_2$ y la estatura media del primer grupo es mayor que la del segundo

Con base en la hipótesis H_0,

$$\mu_{\bar{X}_1 - \bar{X}_2} = 0 \qquad \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(2.5)^2}{50} + \frac{(2.8)^2}{50}} = 0.53$$

donde hemos usado las desviaciones estándar muestrales como estimaciones de σ_1 y σ_2,

Entonces,

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{68.2 - 67.5}{0.53} = 1.32$$

De acuerdo con la prueba de una cola a un nivel de significancia de 0.05, se debería rechazar la hipótesis H_0 si el puntaje z fuera mayor que 1.645. Por tanto, no se puede rechazar la hipótesis a este nivel de significancia.

No obstante, cabe advertir que es posible rechazar la hipótesis a un nivel de 0.10 si se está dispuesto a tomar el riesgo de equivocarse con una probabilidad de 0.10, esto es, 1 posibilidad en 10.

b) El valor P de la prueba es $P(Z \geq 1.32) = 0.0934$, lo cual es la probabilidad de que la diferencia positiva observada entre las medias de las estaturas de atletas hombres y otros estudiantes ocurriría por casualidad si H_0 fuera verdadera.

7.12. ¿Cuánto se tiene que aumentar el tamaño de la muestra de cada uno de los dos grupos del problema 7.11 para que la diferencia observada de 0.7 pulgadas en las medias de las estaturas sea significativa a un nivel de significancia de *a*) 0.05, *b*) 0.01?

Supongamos que el tamaño de la muestra de cada grupo es n y que las desviaciones estándar de los dos grupos se mantienen iguales. Con base en la hipótesis H_0 se tiene que $\mu_{\bar{X}_1 - \bar{X}_2} = 0$ y

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} = \sqrt{\frac{(2.5)^2 + (2.8)^2}{n}} = \sqrt{\frac{14.09}{n}} = \frac{3.75}{\sqrt{n}}$$

Para una diferencia observada de 0.7 pulgadas en las medias de las estaturas

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{0.7}{3.75/\sqrt{n}} = \frac{0.7\sqrt{n}}{3.75}$$

$$\frac{\sqrt{n}}{3.75} \geq 1.645 \quad \text{o bien} \quad \sqrt{n} \geq 8.8 \quad \text{o bien} \quad n \geq 78$$

Por tanto, se debe aumentar el tamaño de la muestra en cada grupo por lo menos $78 - 50 = 28$.

b) La diferencia observada será significativa a un nivel de 0.01 si

$$\frac{0.7\sqrt{n}}{3.75} \geq 2.33 \quad \text{o bien} \quad \sqrt{n} \geq 12.5 \quad \text{o bien} \quad n \geq 157$$

En consecuencia, se debe aumentar el tamaño de la muestra en cada grupo por lo menos $157 - 50 = 107$.

7.13. Dos grupos, *A* y *B*, constan cada uno de 100 personas que tienen una enfermedad. El suero se da al grupo *A* pero no al grupo *B* (que se llama *grupo de control*); en otro caso, los grupos se tratan de manera idéntica. Se encuentra que en los grupos *A* y *B*, con 75 y 65 personas, respectivamente, se recuperan de la enfermedad. Probar la hipótesis de que el suero cura la enfermedad con base en un nivel de significancia de *a*) 0.01, *b*) 0.05, *c*) 0.10. *d*) Calcular el valor *P* de la prueba.

Sean p_1 y p_2, respectivamente, las proporciones de población que se cura 1) usando el suero y 2) sin usarlo. Se tiene que decidir entre dos hipótesis:

H_0: $p_1 = p_2$ y las diferencias observadas son debidas al azar, es decir, el suero no es efectivo

H_1: $p_1 > p_2$ y el suero es efectivo

Con base en la hipótesis H_0

$$\mu_{P_1-P_2} = 0 \qquad \sigma_{P_1-P_2} = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(0.70)(0.30)\left(\frac{1}{100} + \frac{1}{100}\right)} = 0.0648$$

donde se ha usado como una estimación de p la proporción promedio de las curas en los dos grupos muestrales, dada por $(75 + 65)/200 = 0.70$, y donde $q = 1 - p = 0.30$. En consecuencia,

$$Z = \frac{P_1 - P_2}{\sigma_{P_1-P_2}} = \frac{0.750 - 0.650}{0.0648} = 1.54$$

a) Con base en la prueba de una cola a un nivel de significancia de 0.01, se rechaza la hipótesis H_0 sólo si el puntaje z fuera mayor que 2.33. Puesto que el puntaje z es 1.54, concluimos que los resultados se deben a la casualidad en este nivel de significancia.

b) Con base en la prueba de una cola a un nivel de significancia de 0.05, se rechaza la hipótesis H_0 sólo si el puntaje z fuera mayor que 1.645, por lo que se tiene que concluir que los resultados a este nivel también se deben a la casualidad.

c) Si se usara una prueba de una cola a un nivel de significancia de 0.10, se debe rechazar H_0 sólo si el puntaje z fuera mayor que 1.28. Puesto que se satisface esta condición, es posible concluir que el suero es eficaz a un nivel de significancia de 0.10.

d) El valor *P* de la prueba es $P(Z \geq 1.54) = 0.0618$, que es la probabilidad de que ocurriría por casualidad el puntaje z de 1.54 o mayor a favor del grupo de usuarios si H_0 fuera verdadera.

Observamos que las conclusiones anteriores dependen del grado en que se está dispuesto a arriesgar estar equivocado. Si los resultados de hecho se deben a la casualidad y se concluye que se deben al suero (error del tipo I), quizá se proceda a dar el suero a grupos grandes de personas, sólo para descubrir que en realidad es ineficaz. Éste es un riesgo que no siempre se quiere asumir.

Por otro lado, se puede concluir que el suero no ayuda cuando en la realidad sí lo hace (error del tipo II). Esta conclusión es muy peligrosa, sobre todo si están en riesgo vidas humanas.

7.14. Resolver el problema 7.13 si cada grupo consta de 300 personas y si se curan 225 personas en el grupo *A* y 195 en el grupo *B*.

$$\mu_{P_1-P_2} = 0 \qquad \sigma_{P_1-P_2} = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(0.70)(0.30)\left(\frac{1}{300} + \frac{1}{300}\right)} = 0.0374$$

donde se usa $(225 + 195)/600 = 0.70$ como una estimación de p. Entonces,

$$Z = \frac{P_1 - P_2}{\sigma_{P_1-P_2}} = \frac{0.750 - 0.650}{0.0374} = 2.67$$

Puesto que este valor de z es mayor que 2.33, se rechaza la hipótesis a un nivel de significancia de 0.01, esto es, se concluye que el suero es eficaz sólo con una probabilidad de 0.01 de equivocarse. Aquí, el valor P de la prueba es $P(Z \geq 2.67) = 0.0038$.

Este resultado muestra que al aumentar el tamaño de la muestra es posible aumentar la confiabilidad de las decisiones. No obstante, muchas veces es posible que no sea práctico aumentar los tamaños de las muestras. En estos casos resulta obligado tomar decisiones con base en la información disponible, por lo que se enfrenta un mayor riesgo de decisiones equivocadas.

7.15. Una encuesta muestral de 300 votantes del distrito A y 200 votantes del distrito B indicó que 56 y 48%, respectivamente, estaban a favor de cierto candidato. A un nivel de significancia de 0.05, probar la hipótesis de que *a*) existe una diferencia entre los distritos, *b*) se prefiere al candidato en el distrito A. *c*) Calcular los respectivos valores P de la prueba.

Sea p_1 y p_2 la proporción de todos los votantes de los distritos A y B, respectivamente, que están a favor del candidato.

Con base en la hipótesis H_0: $p_1 = p_2$, tenemos

$$\mu_{P_1-P_2} = 0 \qquad \sigma_{P_1-P_2} = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{(0.528)(0.472)\left(\frac{1}{300} + \frac{1}{200}\right)} = 0.0456$$

donde usamos como estimaciones de p y q los valores $[(0.56)(300) + (0.48)(200)]/500 = 0.528$ y $1 - 0.528 = 0.472$. Entonces,

$$Z = \frac{P_1 - P_2}{\sigma_{P_1-P_2}} = \frac{0.560 - 0.480}{0.0456} = 1.75$$

a) Si sólo se desea determinar si hay una diferencia entre los distritos, se debe decidir entre la hipótesis H_0: $p_1 = p_2$ y H_1: $p_1 \neq p_2$, que implica una prueba de dos colas.

Con base en la prueba de dos colas a un nivel de significancia de 0.05, se debería rechazar H_0 si Z estuviera fuera del intervalo de -1.96 a 1.96. Puesto que $Z = 1.75$ yace dentro de este intervalo, no se puede rechazar H_0 a este nivel, esto es, no hay una diferencia significativa entre los distritos.

b) Si se desea determinar si el distrito A prefiere al candidato, se tiene que decidir entre las hipótesis H_0: $p_1 = p_2$ y H_0: $p_1 > p_2$, lo que implica una prueba de una cola.

Con base en la prueba de una cola a un nivel de significancia de 0.05, se debería rechazar H_0 si Z fuera mayor que 1.645. Puesto que éste es el caso, se rechaza H_0 a este nivel y se concluye que el distrito A prefiere al candidato.

c) En la parte *a*), el valor P es $P(Z \leq -1.75) + P(Z \geq 1.75) = 0.0802$, y el valor P en la parte *b*) corresponde a $P(Z \geq 1.75) = 0.0401$.

PRUEBAS EN LAS QUE INTERVIENE LA DISTRIBUCIÓN *T* DE STUDENT

7.16. En el pasado, una máquina producía empaques con un espesor medio de 0.50 pulgadas. Para determinar si la máquina funciona de manera correcta, se elige una muestra de 10 empaques para los cuales el espesor medio es de 0.053 pulgadas y la desviación estándar de 0.003 pulgadas. Probar la hipótesis de que la máquina funciona bien con base en un nivel de significancia de *a*) 0.05, *b*) 0.01. *c*) Encontrar el valor P de la prueba.

$$\mu = 0.050$$
$$\mu \neq 0.050$$

por lo que se necesita una prueba de dos colas.

De acuerdo con la hipótesis H_0, se tiene $T = \dfrac{\bar{X} - \mu}{S}\sqrt{n - 1} = \dfrac{0.053 - 0.050}{0.003}\sqrt{10 - 1} = 3.00$.

a) En el caso de una prueba de dos colas a un nivel de significancia de 0.05, empleamos la regla de decisión:

1) Aceptar H_0 si T yace dentro del intervalo $-t_{0.975}$ a $t_{0.975}$, que para $10 - 1 = 9$ grados de libertad es el intervalo de -2.26 a 2.26.
2) Si no, rechazar H_0.

Puesto que $T = 3.00$, se rechaza H_0 a un nivel de significancia de 0.05.

b) En el caso de una prueba de dos colas a un nivel de significancia de 0.01, utilizamos la regla de decisión:

1) Aceptar H_0 si T yace dentro del intervalo $-t_{0.995}$ a $t_{0.995}$, que para $10 - 1 = 9$ grados de libertad es el intervalo de -3.25 a 3.25.
2) Si no, rechazar H_0.

Debido a que $T = 3.00$, se acepta H_0 a un nivel de significancia de 0.01.

Puesto que es posible rechazar H_0 a un nivel de 0.05 pero no a uno de 0.01, es posible decir que el resultado de la muestra es *probablemente significativo* (vea la terminología al final del problema 7.5). Por tanto, sería recomendable revisar la máquina o al menos tomar otra muestra.

c) El valor P es $P(T \geq 3) + P(T \leq -3)$. La tabla del apéndice D indica que $0.01 < P < 0.02$. Utilizando software de computadora, se encuentra $P = 0.015$.

7.17. Una prueba de la resistencia a la ruptura de 6 cuerdas que fabrica una compañía mostró una resistencia a la ruptura media de 7 750 libras y una desviación estándar de 145 libras, mientras que el fabricante afirma una resistencia a la ruptura de 8 000 libras. ¿Es posible respaldar la afirmación del fabricante a un nivel de significancia de *a*) 0.05, *b*) 0.01? *c*) ¿Cuál es el valor P de la prueba?

Se necesita decidir entre las hipótesis

H_0: $\mu = 8\,000$ lb y la afirmación del fabricante está justificada

H_1: $\mu < 8\,000$ lb y la afirmación del fabricante es falsa

por tanto, se usa una prueba de una cola.

Con base en la hipótesis H_0, se tiene

$$T = \frac{\bar{X} - \mu}{S}\sqrt{n - 1} = \frac{7\,750 - 8\,000}{145}\sqrt{6 - 1} = -3.86.$$

a) En el caso de una prueba de una cola a un nivel de significancia de 0.05, utilizamos la regla de decisión:

1) Aceptar H_0 si T es mayor que $-t_{0.95}$, lo cual con $6 - 1 = 5$ grados de libertad significa $T > -2.01$.
2) Rechazar H_0 en otro caso.

Puesto que $T = -3.86$, se rechaza H_0.

b) En el caso de una prueba de una cola a un nivel de significancia de 0.01, adoptamos la regla de decisión:

1) Aceptar H_0 si T es mayor que $-t_{0.99}$, lo cual para 5 grados de libertad significa $T > -3.36$.
2) Rechazar H_0 en otro caso.

Puesto que $T = -3.86$, se rechaza H_0.
Concluimos que es en extremo improbable que se justifique la afirmación del fabricante.

c) El valor P es $P(T \leq -3.86)$. La tabla del apéndice D muestra que $0.005 < P < 0.01$. Mediante software de computadora, $P = 0.006$.

$$\mu_1 = \mu_2$$
$$\mu_1 \neq \mu_2$$

Con base en la hipótesis H_0,

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{1/n_1 + 1/n_2}} \qquad \text{donde} \qquad \sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}$$

Entonces,

$$\sigma = \sqrt{\frac{16(10)^2 + 14(8)^2}{16 + 14 - 2}} = 9.44 \qquad \text{y} \qquad T = \frac{112 - 107}{9.44\sqrt{1/16 + 1/14}} = 1.45$$

a) Con base en la prueba de dos colas a un nivel de significancia de 0.01, se rechazaría H_0 si T estuviera fuera del rango de $-t_{0.995}$ a $t_{0.995}$, que corresponde al de -2.76 a 2.76 para $n_1 + n_2 - 2 = 16 + 14 - 2 = 28$ grados de libertad.

En consecuencia, no podemos rechazar H_0 a un nivel de significancia de 0.01.

b) Con base en la prueba de dos colas a un nivel de significancia de 0.05, se debería rechazar H_0 si T estuviera fuera del rango de $-t_{0.975}$ a $t_{0.975}$, que para 28 grados de libertad es el rango de -2.05 a 2.05.

Por tanto, no podemos rechazar H_0 a un nivel de significancia de 0.05. Se concluye que no hay diferencia significativa entre los CI de los dos grupos.

c) El valor P es $P(T \geq 1.45) + P(T \leq -1.45)$. La tabla en el apéndice D muestra que $0.1 < P < 0.2$. Mediante software de computadora, $P = 0.158$.

7.19. En un área agrícola, se desea probar el efecto de un fertilizante en la producción de trigo. Para lograr este objetivo, se eligieron 24 lotes de terreno con áreas iguales; se trató una mitad con el fertilizante mientras que la otra mitad no lo recibió (grupo de control). Por lo demás, las condiciones fueron iguales. La producción media de trigo de los lotes sin tratamiento fue de 4.8 quintales, con una desviación estándar de 0.40 quintales, mientras que los lotes tratados produjeron una media de 5.1 quintales con una desviación estándar de 0.36 quintales. ¿Es posible concluir que hay una mejora significativa en la producción de trigo a raíz del fertilizante si se usa un nivel de significancia de *a*) 1%, *b*) 5%? *c*) ¿Cuál es el valor P de la prueba?

Si μ_1 y μ_2 denotan la media poblacional de producción de los lotes de los terrenos tratados y los nos tratados, respectivamente, se tiene que decidir entre las hipótesis

H_0: $\mu_1 = \mu_2$ y la diferencia se debe al azar

H_1: $\mu_1 > \mu_2$ y el fertilizante incrementa la producción

Con base en la hipótesis H_0,

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{1/n_1 + 1/n_2}} \qquad \text{donde} \qquad \sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}$$

Entonces,

$$\sigma = \sqrt{\frac{12(0.40)^2 + 12(0.36)^2}{12 + 12 - 2}} = 0.397 \qquad \text{y} \qquad T = \frac{5.1 - 4.8}{0.397\sqrt{1/12 + 1/12}} = 1.85$$

PRUEBAS EN LAS QUE INTERVIENE LA DISTRIBUCIÓN JI-CUADRADA

7.20. En el pasado, la desviación estándar de los pesos de ciertos paquetes de 40.0 onzas que se llenaban mediante una máquina era de 0.25 onzas. Una muestra aleatoria de 20 paquetes mostró una desviación estándar de 0.32 onzas. ¿Es el aumento en la variabilidad significativo a un nivel de significancia de *a*) 0.05, *b*) 0.01? *c*) ¿Cuál es el valor *P* de la prueba?

Se necesita decidir entre las hipótesis

H_0: $\sigma = 0.25$ onzas y los resultados observados se deben al azar

H_1: $\sigma > 0.25$ onzas y la variabilidad se ha incrementado

El valor de χ^2 de la muestra es $\chi^2 = ns^2/\sigma^2 = 20(0.32)^2/(0.25)^2 = 32.8$.

a) Usando una prueba de una cola, se rechazaría H_0 a un nivel de significancia de 0.05 si el valor muestral de χ^2 fuera mayor que $\chi^2_{0.95}$, que equivale a 30.1 para $\nu = 20 - 1 = 19$ grados de libertad. En consecuencia, se rechaza H_0 a un nivel de significancia de 0.05.

b) Con base en una prueba de una cola, se debería rechazar H_0 a un nivel de significancia de 0.01 si el valor muestral de χ^2 fuera mayor que $\chi^2_{0.99}$, que equivale a 36.2 para 19 grados de libertad. Por tanto, se debería rechazar H_0 a un nivel de significancia de 0.01.

Concluimos que la variabilidad tiene una mayor probabilidad. Sería necesario revisar la máquina.

c) El valor *P* es $P(\chi^2 \geq 32.8)$. La tabla del apéndice E muestra que $0.025 < P < 0.05$. Mediante software de computadora, $P = 0.0253$.

PRUEBAS EN LAS QUE INTERVIENE LA DISTRIBUCIÓN *F*

7.21. Un instructor atiende a dos grupos, *A* y *B*, en un tema particular. La clase *A* tiene 16 estudiantes mientras que la *B* tiene 25. En el mismo examen, aunque no hubo diferencias significativas entre las medias de las calificaciones, la clase *A* mostró una desviación estándar de 9, mientras que la de la clase *B* fue de 12. ¿Es posible concluir que la variabilidad de la clase *B* es mayor que la de la clase *A* a un nivel de significancia *a*) 0.01, *b*) 0.05? *c*) ¿Cuál es el valor *P* de la prueba?

a) Al usar los subíndices 1 y 2 para las clases *A* y *B*, respectivamente, se tiene $s_1 = 9$, $s_2 = 12$, por lo que

$$\hat{s}_1^2 = \frac{n_1}{n_1 - 1} s_1^2 = \frac{16}{15}(9)^2 = 86.4, \qquad \hat{s}_2^2 = \frac{n_2}{n_2 - 1} s_2^2 = \frac{25}{24}(12)^2 = 150$$

Se tiene que decidir entre las hipótesis

H_0: $\sigma_1 = \sigma_2$ y cualquier variabilidad observada se debe al azar

H_1: $\sigma_2 > \sigma_1$ y la variabilidad de la clase *B* es mayor que la de *A*

Por tanto, la decisión tiene que basarse en una prueba de una cola de la distribución *F*. En el caso de las muestras en cuestión,

$$F = \frac{\hat{s}_2^2}{\hat{s}_1^2} = \frac{150}{86.4} = 1.74$$

CURVAS CARACTERÍSTICAS DE OPERACIÓN

7.23. Con referencia al problema 7.2, ¿cuál es la probabilidad de aceptar la hipótesis de que la moneda no está cargada cuando la probabilidad real de caras es $p = 0.7$?

La hipótesis H_0 de que la moneda no está cargada, esto es, $p = 0.5$, se acepta cuando el número de caras en 100 lanzamientos yace entre 39.5 y 60.5. La probabilidad de rechazar H_0 cuando se debe aceptar (esto es, la probabilidad de un error del tipo I) se representa por el área total α de la región sombreada a la izquierda bajo la curva normal de la figura 7-6. Como se calcula en el problema 7.2 *a*), esta área α, que representa el nivel de significancia de la prueba H_0 es igual a 0.0358.

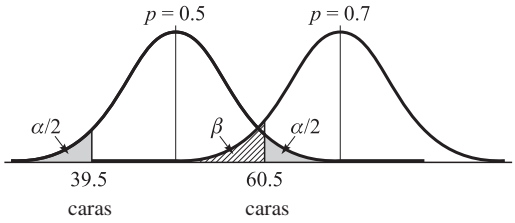


Figura 7-6

Si la probabilidad de caras es $p = 0.7$, entonces se representa la distribución de caras en 100 lanzamientos mediante la curva normal a la derecha de la figura 7-6. La probabilidad de aceptar H_0 cuando en efecto $p = 0.7$ (esto es, la probabilidad de un error del tipo II) está dada por el área rayada β. Para calcular esta área, se observa que la distribución con base en la hipótesis $p = 0.7$ tiene una media y una desviación estándar dadas por

$$\mu = np = (100)(0.7) = 70 \qquad \sigma = \sqrt{npq} = \sqrt{(100)(0.7)(0.3)} = 4.58$$

$$60.5 \text{ en unidades estándar} = \frac{60.5 - 70}{4.58} = -2.07$$

$$39.5 \text{ en unidades estándar} = \frac{39.5 - 70}{4.58} = -6.66$$

En consecuencia, β = área bajo la curva normal estándar entre $z = -6.66$ y $z = -2.07 = 0.0192$.

Por tanto, con la regla de decisión dada hay poca posibilidad de aceptar la hipótesis de que la moneda no está cargada cuando efectivamente $p = 0.7$.

Observemos que en este problema se dio la regla de decisión a partir de la cual se calcularon α y β. En la práctica, pueden surgir otras dos posibilidades:

$$\mu = np = (100)(0.6) = 60 \qquad \sigma = \sqrt{npq} = \sqrt{(100)(0.6)(0.4)} = 4.90$$

$$60.5 \text{ en unidades estándar} = \frac{60.5 - 60}{4.90} = 0.102$$

$$39.5 \text{ en unidades estándar} = \frac{39.5 - 60}{4.90} = -4.18$$

En consecuencia, β = área bajo la curva normal estándar entre $z = -4.18$ y $z = 0.102 = 0.5405$.

Por tanto, con la regla de decisión dada existe una gran posibilidad de aceptar la hipótesis de que la moneda no está cargada cuando efectivamente $p = 0.6$.

b) Si $p = 0.8$, entonces $\mu = np = (100)(0.8) = 80$ y $\sigma = \sqrt{npq} = \sqrt{(100)(0.08)(0.2)} = 4$.

$$60.5 \text{ en unidades estándar} = \frac{60.5 - 80}{4} = -4.88$$

$$39.5 \text{ en unidades estándar} = \frac{39.5 - 80}{4} = -10.12$$

Entonces, β = área bajo la curva normal estándar entre $z = -10.12$ y $z = -4.88 = 0.0000$, muy aproximadamente.

c) Mediante comparación con *b*) o mediante cálculo, se advierte que si $p = 0.9$, $\beta = 0$ para propósitos prácticos.

d) Por simetría, $p = 0.4$ produce el mismo valor de β como $p = 0.6$, esto es, $\beta = 0.5405$.

7.25. Representar los resultados de los problemas 7.23 y 7.24 mediante una gráfica de *a*) β *vs*. p, *b*) $(1 - \beta)$ *vs*. p. Interpretar las gráficas que se obtengan.

La tabla 7-3 muestra los valores de β correspondientes a valores dados de p según se obtuvieron en los problemas 7.23 y 7.24.

Observemos que β representa la probabilidad de aceptar la hipótesis $p = 0.5$ cuando p es en realidad un valor diferente a 0.5. Sin embargo, si es verdad que $p = 0.5$, se puede interpretar β como la probabilidad de aceptar $p = 0.5$ cuando éste debe aceptarse. Esta probabilidad es igual a $1 - 0.0358 = 0.9642$ y se ha incluido en la tabla 7-3.

Tabla 7-3

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
β	0.0000	0.0000	0.0192	0.5405	0.9642	0.5405	0.0192	0.0000	0.0000

a) La gráfica de β *vs*. p que se muestra en la figura 7-7*a*) se llama *curva característica de operación*, o *curvas CO*, de la regla de decisión o de la prueba de hipótesis.

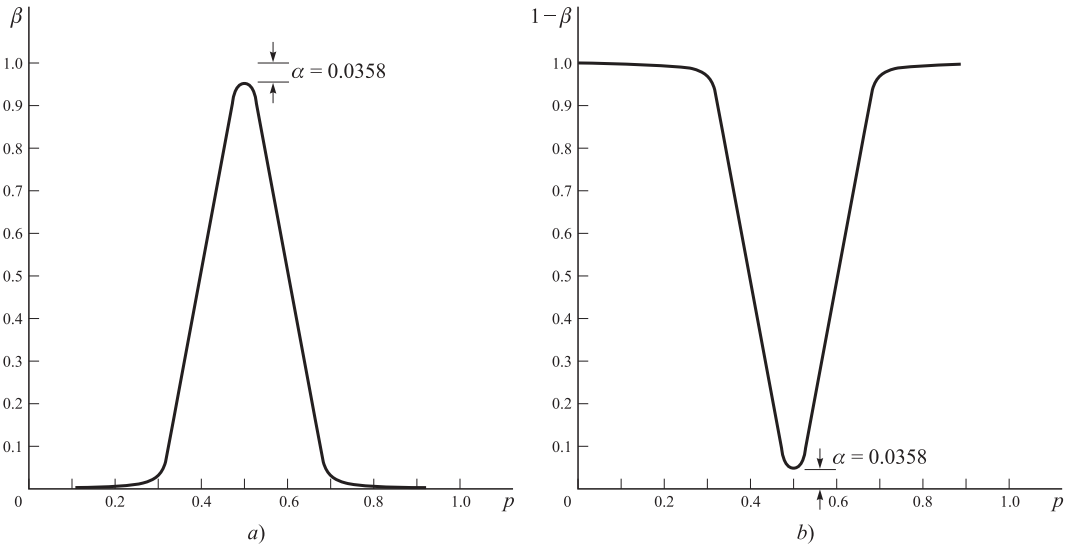
La distancia del punto máximo de la curva CO a la línea $\beta = 1$ es igual a $\alpha = 0.0358$, el nivel de significancia de la prueba.

Por lo general, cuanto más agudo sea el pico de la curva CO, tanto mejor es la regla de decisión para rechazar hipótesis que no son válidas.

b) La gráfica de $(1 - \beta)$ *vs*. p, que se muestra en la figura 7-7*b*), se denomina *curva de potencia* de la regla de decisión o de la prueba de hipótesis. Esta curva se obtiene si simplemente se invierte la curva CO, por lo que en realidad las dos gráficas son equivalentes.

La cantidad $(1 - \beta)$ se conoce a menudo como *función potencia*, puesto que indica la capacidad o *potencia de una prueba* de rechazar hipótesis que son falsas, esto es, que deben rechazarse. La cantidad β también se denomina la *función característica de operación* de una prueba.

7.26. Una compañía fabrica cuerda cuyas resistencias a la ruptura tienen media de 300 libras y desviación estándar de 24 libras. Se cree que mediante un nuevo proceso es posible aumentar la media de la resistencia a la



$$\mu = 300$$

$$\mu > 300$$

En el caso de una prueba de una cola a un nivel de significancia de 0.01, se tiene la siguiente regla de decisión (vea la figura 7-8):

1) Rechazar H_0 si el puntaje z de la media de la resistencia a la ruptura de la muestra es mayor que 2.33.
2) Aceptar la hipótesis H_0 en otro caso.

Puesto que $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \dfrac{\bar{X} - 300}{24/\sqrt{64}}$, $\bar{X} = 300 + 3z$. Entonces si $Z > 2.33$, $\bar{X} > 300 + 3(2.33) = 307.0$ libras.

Por tanto, la regla de decisión anterior se convierte en:

1) Rechazar H_0 si la media de la resistencia a la ruptura de 64 cuerdas es mayor que 307.0 libras.
2) Aceptar H_0 en otro caso.

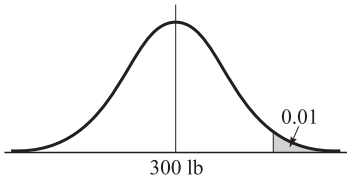


Figura 7-8

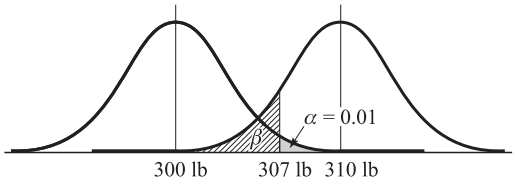


Figura 7-9

b) Consideremos las dos hipótesis (H_0:$\mu = 300$ libras) y (H_1:$\mu = 310$ libras). Las distribuciones de las resistencias a la ruptura correspondientes a estas dos hipótesis están representadas, respectivamente, por las distribuciones normales de la izquierda y la derecha de la figura 7-9.

La probabilidad de aceptar el viejo proceso cuando la nueva resistencia a la ruptura media es en realidad de 310 libras está representada por la región del área β en la figura 7-9. Para determinar esta probabilidad, observemos que 307.0 libras en unidades estándar es $(307.0 - 310)/3 = -1.00$; por tanto,

$$\beta = \text{área bajo la curva normal de la derecha a la izquierda de } z = -1.00 = 0.1587$$

Tabla 7-4

μ							
β							

a) La curva CO se muestra en la figura 7-10*a*). A partir de esta curva se observa que la probabilidad de mantener el viejo proceso si la nueva resistencia a la ruptura es menor que 300 libras es prácticamente 1 (excepto para el nivel de significancia de 0.01, cuando el nuevo proceso resulta en una media de 300 libras). Después ésta cae bastante rápido a cero, por lo que prácticamente no hay ninguna posibilidad de mantener el proceso primitivo cuando la resistencia a la ruptura media es mayor que 315 libras.

b) La curva de potencia que se ilustra en la figura 7-10*b*) la interpretamos exactamente de la misma manera que la correspondiente a la curva CO. En realidad, las dos curvas son, en esencia, equivalentes.

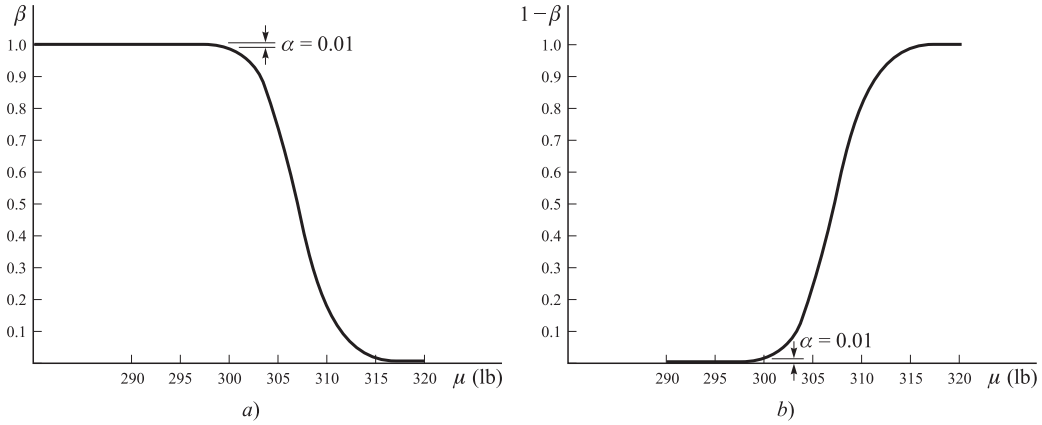


Figura 7-10

7.28. Para probar la hipótesis de que una moneda no está cargada (esto es, $p = 0.5$) mediante un número de lanzamientos, se imponen las siguientes restricciones: *a*) la probabilidad de rechazar la hipótesis cuando la moneda es legal debe ser a lo más 0.05; *b*) la probabilidad de aceptar la hipótesis cuando p difiere de 0.5 por 0.1 o más (esto es, $p \geq 0.6$ o bien $p \leq 0.4$) debe ser cuando mucho 0.05. Determinar el tamaño de muestra mínimo que es necesario y enunciar la regla de decisión resultante.

Aquí se han puesto límites a los riesgos de los errores del tipo I y del tipo II. Por ejemplo, la restricción *a*) requiere que la probabilidad de un error del tipo I sea de $\alpha = 0.05$ como máximo, mientras que la restricción *b*) requiere que la probabilidad de error máximo del tipo II sea $\beta = 0.05$. La figura 7-11 ilustra la situación.

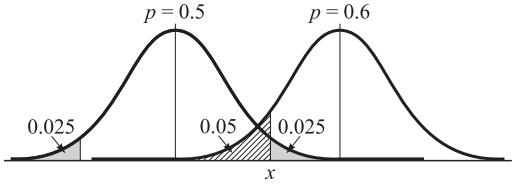


Figura 7-11

$$\frac{-}{\sqrt{npq}} = \frac{x - 0.5n}{0.5\sqrt{n}}$$

2) El área bajo la curva normal con $p = 0.6$ a la izquierda de $\dfrac{x - np}{\sqrt{npq}} = \dfrac{x - 0.6n}{0.49\sqrt{n}}$ es 0.05.

En realidad, se debió haber igualado el área entre

$$\frac{(n - x) - 0.6n}{0.49\sqrt{n}} \quad \text{y} \quad \frac{x - 0.6n}{0.49\sqrt{n}}$$

a 0.05; no obstante, 2) es una buena aproximación. Observe que al hacer la probabilidad de aceptación igual a 0.05 en "el peor caso", $p = 0.6$, ésta se convierte automáticamente en 0.05 o menor cuando p tiene cualquier otro valor fuera del rango de 0.4 a 0.6. En consecuencia, un promedio ponderado de todas estas probabilidades, lo cual representa la probabilidad de un error del tipo II, será también de 0.05 o menor.

De 1), $\dfrac{x - 0.5n}{0.5\sqrt{n}} = 1.96$ o bien 3) $x = 0.5n + 0.980\sqrt{n}$.

De 2), $\dfrac{x - 0.6n}{0.49\sqrt{n}} = -1.645$ o bien 4) $x = 0.6n - 0.806\sqrt{n}$.

En consecuencia, de 3) y 4), $n = 318.98$. Se concluye que el tamaño de la muestra tiene que ser por lo menos de 319, esto es, se necesita lanzar la moneda un mínimo de 319 veces. Cuando se sustituye $n = 319$ en 3) o en 4), $x = 177$.

Para $p = 0.5$, $x - np = 177 - 159.5 = 17.5$. Por tanto, seguimos la siguiente regla de decisión:

a) Aceptar la hipótesis $p = 0.5$ si el número de caras en 319 lanzamientos está en el rango de 159.5 ± 17.5, esto es, entre 142 y 177.

b) Rechazar la hipótesis en otro caso.

GRÁFICAS DE CONTROL DE CALIDAD

7.29 Se construye una máquina para producir cojinetes de bolas cuya media de diámetro es 0.574 pulgadas y su desviación estándar es de 0.008 pulgadas. Para determinar si la máquina funciona de manera adecuada, cada 2 horas se toma una muestra de 6 cojinetes de bolas y se calcula la media del diámetro de ella. *a*) Diseñar una regla de decisión que otorgue bastante seguridad de que la calidad de los productos es acorde a los estándares requeridos. *b*) Demostrar cómo representar gráficamente la regla de decisión de *a*).

a) Con 99.73% de confianza se puede decir que la media muestral $\overline{X}$ debe ubicarse en el rango $(\mu_{\bar{X}} - 3\sigma_{\bar{X}})$ a $(\mu_{\bar{X}} + 3\sigma_{\bar{X}})$ o bien $(\mu - 3\sigma/\sqrt{n})$ a $(\mu + 3\sigma/\sqrt{n})$. Puesto que $\mu = 0.574$, $\sigma = 0.008$ y $n = 6$, se deduce que con una confianza de 99.73% la media muestral debe estar entre $(0.574 - 0.024/\sqrt{6})$ y $(0.574 + 0.024/\sqrt{6})$ o entre 0.564 y 0.584 pulgadas.

En consecuencia, la regla de decisión es la siguiente:

1) Si la media muestral está dentro del rango 0.564 a 0.584 pulgadas, se debe suponer que la máquina funciona de manera adecuada.

2) En otro caso, es necesario concluir que la máquina no funciona bien y se debe determinar la causa.

b) Se puede mantener un registro de las medias muestrales mediante una gráfica como la que se muestra en la figura 7-12, denominada *gráfica de control de calidad*. Cada vez que se calcula una media muestral, ésta se representa por medio de un punto. Siempre que los puntos se encuentren entre el límite inferior de 0.564 pulgadas y el límite superior de 0.584 pulgadas, el proceso está controlado. Cuando un punto queda fuera de estos límites de control (como la tercera muestra tomada el jueves), existe la posibilidad de que algo incorrecto esté pasando y merece una investigación.

Los límites de control especificados antes se llaman límites de confianza de 99.73%, o con brevedad, los límites de 3σ. Sin embargo, se pueden determinar otros límites de confianza, como los de 99 y 95%. En cada caso, la selección depende de las circunstancias particulares.

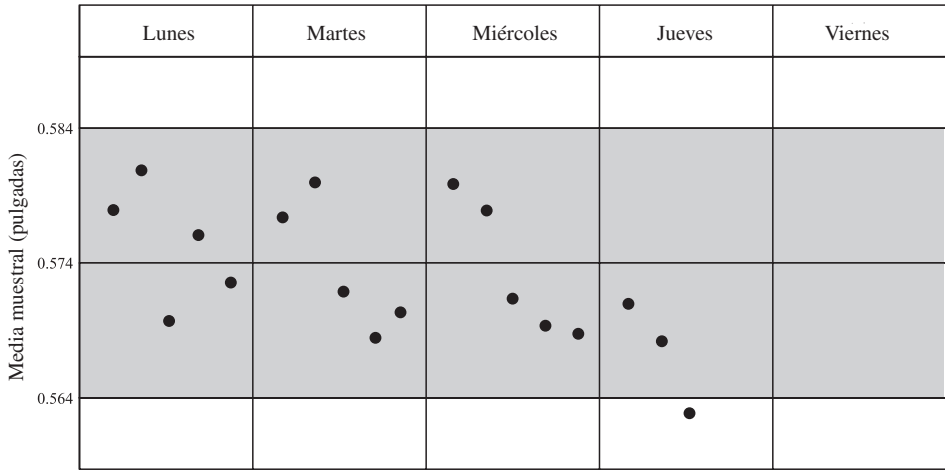


Figura 7-12

AJUSTE DE DATOS MEDIANTE DISTRIBUCIONES TEÓRICAS

7.30. Ajustar una distribución binomial a los datos del problema 5.30, página 176.

Se tiene $P(x \text{ caras en un lanzamiento de 5 monedas}) = f(x) = {}_5C_x p^x q^{5-x}$, donde p y q son las probabilidades respectivas de una cara y una cruz en un solo lanzamiento de la moneda. La media o número esperado de caras es $\mu = np = 5p$.

Para la distribución de frecuencias verdadera u observada, la media del número de caras es

$$\frac{\sum fx}{\sum f} = \frac{(38)(0) + (144)(1) + (342)(2) + (287)(3) + (164)(4) + (25)(5)}{1\,000} = \frac{2\,470}{1\,000} = 2.47$$

Al igualar las medias teórica y verdadera, $5p = 2.47$ o bien $p = 0.494$. En consecuencia, la distribución binomial ajustada está dada por $f(x) = {}_5C_x = (0.494)^x(0.506)^{5-x}$.

En la tabla 7-5 se ha hecho una lista de estas probabilidades además de las frecuencias esperadas (teóricas) y las reales. Se observa que el ajuste es aceptable. Se investiga la bondad de ajuste en el problema 7.43.

Tabla 7-5

Número de caras (x)	$P(x \text{ caras})$	Frecuencias esperadas	Frecuencias observadas
0	0.0332	33.2 o 33	38
1	0.1619	161.9 o 162	144
2	0.3162	316.2 o 316	342
3	0.3087	308.7 o 309	287
4	0.1507	150.7 o 151	164
5	0.0294	29.4 o 29	25

7.31. Utilizar papel para gráficas de probabilidad con el fin de determinar si es posible aproximar la distribución de frecuencias de la tabla 5-2, página 161, mediante una distribución normal.

Primero se convierte la primera distribución de frecuencias dadas en una distribución de frecuencias relativas acumuladas, como se muestra en la tabla 7-6. Luego, las frecuencias relativas acumuladas expresadas como porcentajes se dibujan contra las fronteras de clase superiores en papel especial para graficar probabilidades, como se muestra en la figura 7-13. El grado al cual todos los puntos dibujados están en una línea recta determina la precisión de ajuste de la distribución dada a una distribución normal. Se advierte que hay una distribución normal que se aproxima mucho a los datos. Vea el problema 7.32.

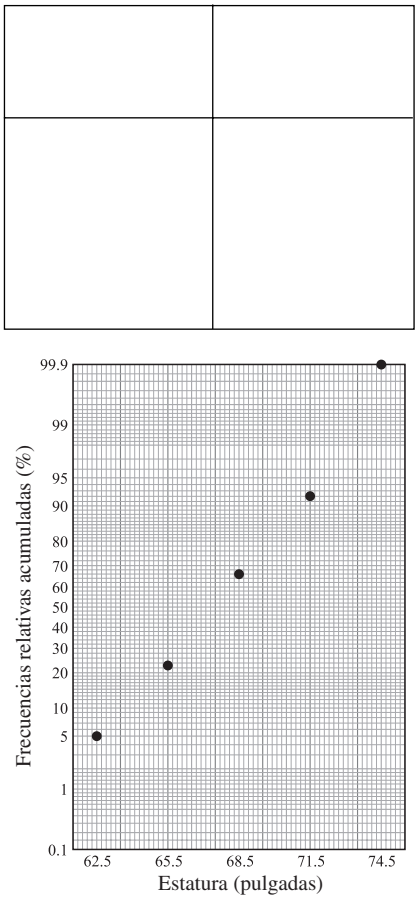


Figura 7-13

7.32. Ajustar una curva normal a los datos de la tabla 5-2, página 161.

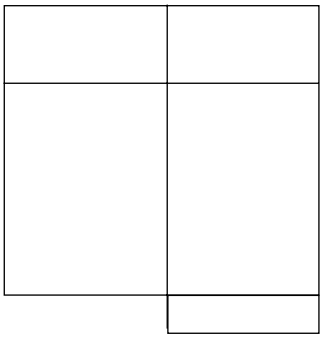
$$\bar{x} = 67.45 \text{ pulgadas}, \qquad s = 2.92 \text{ pulgadas}$$

Es posible organizar la tarea como en la tabla 7-7. Al calcular z para las fronteras de clase, se utiliza $z = (x - \bar{x})/s$ donde se obtuvieron la media $\bar{x}$ y las desviaciones estándar s en los problemas 5.35 y 5.40, respectivamente.

Tabla 7-7

Estaturas (pulgadas)	Límites reales de clase (x)	z para los límites reales de clase	Área bajo la curva normal desde 0 hasta z	Área para cada clase	Frecuencias esperadas	Frecuencias observadas
	59.5	−2.72	0.4967			
60-62				0.0413	4.13 o 4	5
	62.5	−1.70	0.4554			
63-65				0.2086	20.68 o 21	18
	65.5	−0.67	0.2486			
66-68			} Sumar →	0.3892	38.92 o 39	42
	68.5	0.36	0.1406			
69-71				0.2771	27.71 o 28	27
	71.5	1.39	0.4177			
72-74				0.0743	7.43 o 7	8
	74.5	2.41	0.4920			

En la cuarta columna, se obtuvieron las áreas bajo de la curva normal de 0 a z mediante la tabla del apéndice C. De aquí se puede partir para encontrar las áreas debajo de la curva normal entre valores sucesivos de z, como se muestra en la quinta columna. A partir de este punto se obtienen las áreas bajo la curva normal entre valores suce-



$$\lambda = \frac{\sum fx}{\sum f} = \frac{(21)(0) + (18)(1) + (7)(2) + (3)(3) + (1)(4)}{50}$$

$$= \frac{45}{50} = 0.90$$

Entonces, de acuerdo con la distribución de Poisson,

$$P(x \text{ accidentes}) = \frac{(0.90)^x e^{-0.90}}{x!}$$

En la tabla 7-9 se ha hecho una lista de las probabilidades de que ocurran 0, 1, 2, 3 y 4 accidentes que se obtuvieron de esta distribución de Poisson, así como el número teórico de días durante los cuales ocurrieron z accidentes (resultado que se obtiene al multiplicar las probabilidades respectivas por 50). Para conveniencia de la comparación, se repite la cuarta columna que proporciona el número verdadero de días.

Tabla 7-9

Número de accidentes (x)	P (x accidentes)	Número de días esperado	Número de días observado
0	0.4066	20.33 o 20	21
1	0.3659	18.30 o 18	18
2	0.1647	8.24 o 8	7
3	0.0494	2.47 o 2	3
4	0.0111	0.56 o 1	1

Observemos que el ajuste de la distribución de Poisson a los datos es bueno.

En una verdadera distribución de Poisson, $\sigma^2 = \lambda$. El cálculo de la varianza de la distribución dada es 0.97, que se compara de manera favorable con el valor de 0.90 de λ, que puede considerarse como evidencia adicional en cuanto a lo apropiado de la distribución de Poisson para aproximar los datos muestrales.

PRUEBA JI CUADRADA

7.34. En 200 lanzamientos de una moneda, se observaron 115 caras y 85 cruces. Probar la hipótesis de que la moneda no está cargada con base en un nivel de significancia de *a*) 0.05, *b*) 0.01. *c*) Encontrar el valor *P* de la prueba.

Las frecuencias observadas de caras y cruces son $x_1 = 115$, $x_2 = 85$, respectivamente.

Las frecuencias esperadas de caras y cruces si la moneda no está cargada son $np_1 = 100$, $np_2 = 100$, respectivamente. En consecuencia,

$$x^2 = \frac{(x_1 - np_1)^2}{np_1} + \frac{(x_2 - np_2)^2}{np_2} = \frac{(115 - 100)^2}{100} + \frac{(85 - 100)^2}{100} = 4.50$$

Puesto que el número de categorías o clases (caras, cruces) es $k = 2$, $\nu = k - 1 = 2 - 1 = 1$.

a) El valor crítico $\chi^2_{0.95}$ de 1 grado de libertad es 3.84. Puesto que $4.50 > 3.84$, se rechaza la hipótesis de que la moneda no está cargada a un nivel de significancia de 0.05.

b) El valor crítico $\chi^2_{0.99}$ de un grado de libertad es 6.63. Puesto que $4.50 < 6.63$, no se puede rechazar la hipótesis de que la moneda no está cargada a un nivel de significancia de 0.01.

Concluimos que los resultados observados son *probablemente significativos* y que la moneda está *probablemente cargada*. Para comparar este método con otros anteriores, vea el método 1 del problema 7.36.

c) El valor *P* es $P(\chi^2 \geq 4.50)$. La tabla del apéndice E muestra que $0.025 < P < 0.05$. Por medio de software de computadora, $P = 0.039$.

7.35. Resolver el problema 7.34 con base en la corrección de Yates.

$$\chi^2\,(\text{corregida}) = \frac{(|x_1 - 1 - np_1| - 0.5)^2}{np_1} + \frac{(|x_2 - np_2| - 0.5)^2}{np_2}$$

$$= \frac{(|115 - 100| - 0.5)^2}{100} + \frac{(|85 - 100| - 0.5)^2}{100} = \frac{(14.5)^2}{100} + \frac{(14.5)^2}{100} = 4.205$$

El valor corregido de *P* es 0.04.

Puesto que $4.205 > 3.84$ y $4.205 < 6.63$, resultan válidas las conclusiones a las que se llegó en el problema 7.34. Para comparar con métodos anteriores, vea el método 2 del problema 7.36.

7.36. Resolver el problema 7.34 empleando la aproximación normal a la distribución binomial.

Con base en la hipótesis de que la moneda no está cargada, la media y la desviación estándar del número de caras en 200 lanzamientos de una moneda son $\mu = np = (200)(0.5) = 100$ y $\sigma = \sqrt{npq} = \sqrt{(200)(0.5)(0.5)} = 7.07$.

Método 1

$$115 \text{ caras en unidades estándar} = (115 - 100)/7.07 = 2.12.$$

Con base en un nivel de significancia de 0.05 y una prueba de dos colas, se debería rechazar la hipótesis de que la moneda no está cargada si el puntaje *z* estuviera fuera del intervalo de -1.96 a 1.96. Con un nivel de 0.01, el intervalo correspondiente sería de -2.58 a 2.58. Como en el problema 7.34, se concluye que es posible rechazar la hipótesis a un nivel de significancia de 0.05 pero no a uno de 0.01. El valor *P* de la prueba es 0.034.

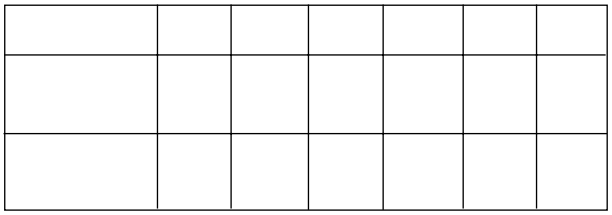
Observemos que el cuadrado del puntaje estándar anterior, $(2.12)^2 = 4.50$, es igual al valor de χ^2 que se obtuvo en el problema 7.34. Éste es siempre el caso para una prueba ji-cuadrada que abarca dos categorías. Vea el problema 7.60.

Método 2

Con base en la corrección de continuidad, 115 o más caras es equivalente a 114.5 caras o más. En consecuencia, 114.5 en unidades estándar $= (114.5 - 100)/7.07 = 2.05$. Esto nos lleva a la misma conclusión que con el método 1. El valor *P* corregido es 0.04.

Observemos que el cuadrado del puntaje estándar anterior corresponde a $(2.05)^2 = 4.20$, lo que concuerda con el valor de χ^2 corregido por continuidad mediante la corrección de Yates del problema 7.35. Éste es siempre el caso de la prueba ji-cuadrada que abarca dos categorías a las cuales se aplica la corrección de Yates, de nuevo, como consecuencia del problema 7.60.

7.37. La tabla 7-10 muestra las frecuencias observadas y esperadas al lanzar un dado 120 veces. *a*) Probar la hipótesis de que el dado no está cargado con el uso de un nivel de significancia de 0.05. *b*) Determinar el valor *P* de la prueba.



$$\chi^2 = \frac{(x_1 - np_1)^2}{np_1} + \frac{(x_2 - np_2)^2}{np_2} + \frac{(x_3 - np_3)^2}{np_3} + \frac{(x_4 - np_4)^2}{np_4} + \frac{(x_5 - np_5)^2}{np_5} + \frac{(x_6 - np_6)^2}{np_6}$$

$$= \frac{(25 - 20)^2}{20} + \frac{(17 - 20)^2}{20} + \frac{(15 - 20)^2}{20} + \frac{(23 - 20)^2}{20} + \frac{(24 - 20)^2}{20} + \frac{(16 - 20)^2}{20} = 5.00$$

Puesto que el número de categorías o clases (caras 1, 2, 3, 4, 5, 6) es $k = 6$, $\nu = k - 1 = 6 - 1 = 5$.

El valor crítico $\chi^2_{0.95}$ de 5 grados de libertad es 11.1. En razón de que $5.00 < 11.1$, no se puede rechazar la hipótesis de que el dado no está cargado.

En el caso de 5 grados de libertad $\chi^2_{0.05} = 1.15$, por lo que $\chi^2 = 5.00 > 1.15$. Se concluye que la concordancia no es excepcionalmente buena de modo que queden dudas.

b) El valor P de la prueba es $P(\chi^2 \geq 5.00)$. La tabla del apéndice E muestra que $0.25 < P < 0.5$. Mediante software de computadora, $P = 0.42$.

7.38. Una tabla de números aleatorios de 250 dígitos tiene la distribución de dígitos, 0, 1, 2, . . . , 9 que se muestra en la tabla 7-11. *a*) ¿La distribución observada difiere de manera significativa de la distribución esperada? *b*) ¿Cuál es el valor P de la observación?

Tabla 7-11

Dígito	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Frecuencias observadas	17	31	29	18	14	20	35	30	20	36
Frecuencias esperadas	25	25	25	25	25	25	25	25	25	25

$$\chi^2 = \frac{(17 - 25)^2}{25} + \frac{(31 - 25)^2}{25} + \frac{(29 - 25)^2}{25} + \frac{(18 - 25)^2}{25} + \cdots + \frac{(36 - 25)^2}{25} = 23.3$$

El valor crítico $\chi^2_{0.99}$ para $\nu = k - 1 = 9$ grados de libertad es 21.7, y $23.3 > 21.7$. Por tanto, se concluye que la distribución observada difiere de manera significativa de la distribución esperada a un nivel de significancia de 0.01, por lo que subsisten dudas con respecto a alguna tendencia no aleatoria en dicha tabla de números.

b) El valor P es $P(\chi^2 = 23.3)$. La tabla del apéndice E muestra que $0.005 < P < 0.01$. Por medio de software de computadora, $P = 0.0056$.

7.39. En sus experimentos con chícharos, Mendel observó 315 redondos y amarillos, 108 redondos y verdes, 101 arrugados y amarillos y 32 arrugados y verdes. De acuerdo con su teoría de la herencia, los números deben estar en la proporción de 9:3:3:1. ¿Existe alguna evidencia para dudar de su teoría a un nivel de significancia de *a*) 0.01, *b*) 0.05? *c*) ¿Cuál es el valor P de la observación?

El número total de chícharos es $315 + 108 + 101 + 32 = 556$. Puesto que los números esperados tienen la proporción 9:3:3:1 (y $9 + 3 + 3 + 1 = 16$), se debería esperar que

$\frac{9}{16}(556) = 312.75$ redondos y amarillos $\quad$ $\frac{3}{16}(556) = 104.25$ arrugados y amarillos

$\frac{3}{16}(556) = 104.25$ redondos y verdes $\quad$ $\frac{1}{16}(556) = 34.75$ arrugados y verdes

$$\chi = \frac{(315 - 312.75)^2}{312.75} + \frac{(108 - 104.25)^2}{104.25} + \frac{(101 - 104.25)^2}{104.25} + \frac{(32 - 34.75)^2}{37.75} = 0.470$$

Puesto que hay cuatro categorías, $k = 4$ y el número de grados de libertad es $\nu = 4 - 1 = 3$.

a) Para $\nu = 3$, $\chi^2_{0.99} = 11.3$, por lo que no se puede rechazar la teoría a un nivel de 0.01.

b) Para $\nu = 3$, $\chi^2_{0.95} = 7.81$, por lo que no es posible rechazar la teoría a un nivel de 0.05.

Concluimos que la teoría y el experimento concuerdan.

Observemos que para 3 grados de libertad, $\chi^2_{0.05} = 0.352$ y $\chi^2 = 0.470 > 0.352$. Por tanto, aunque la concordancia es buena, los resultados obtenidos están sujetos en gran medida a errores de muestreo.

c) El valor P es $P(\chi^2 \geq 0.470)$. La tabla del apéndice E muestra que $0.9 < P < 0.95$. Mediante software de computadora, $P = 0.93$.

7.40. Una urna contiene un número muy grande de canicas de cuatro colores diferentes: rojo, naranja, amarillo y verde. Una muestra aleatoria de 12 canicas de la urna revela 2 canicas rojas, 5 naranja, 4 amarillas y 1 verde. Probar la hipótesis de que la urna contiene proporciones iguales de canicas de diferentes colores y calcular el valor P de los resultados de la muestra.

Con base en la hipótesis de que la urna contiene proporciones iguales de canicas de diferentes colores, se deberían esperar 3 de cada tipo en una muestra de 12 canicas.

Debido a que estos números esperados son menores que 5, la aproximación de ji-cuadrada estará equivocada. Para evitar el error, se combinan categorías para que el número esperado de cada categoría sea por lo menos 5.

Si se desea rechazar la hipótesis, se deberían combinar las categorías para que la evidencia contra la hipótesis se destaque de la mejor manera. Esto se logra en este caso al considerar las categorías "roja o verde" y "naranja o amarilla", para las cuales la muestra reveló 3 y 9 canicas, respectivamente. Puesto que el número esperado en cada categoría con base en la hipótesis de proporciones iguales es 6, se tiene

$$\chi^2 = \frac{(3 - 6)^2}{6} + \frac{(9 - 6)^2}{6} = 3$$

Para $\nu = 2 - 1 = 1$, $\chi^2_{0.95} = 3.84$. En consecuencia, no podemos rechazar la hipótesis a un nivel de significancia de 0.05 (aunque sí es posible al nivel de 0.10). Es concebible que los resultados observados surjan con base en la casualidad aun cuando están presentes iguales proporciones de colores. El valor P es $P(\chi^2 \geq 3) = 0.083$.

Otro método

Utilizando la corrección de Yates, se encuentra

$$\chi^2 = \frac{(|3 - 6| - 0.5)^2}{6} + \frac{(|9 - 6| - 0.5)^2}{6} = \frac{(2.5)^2}{6} + \frac{(2.5)^2}{6} = 2.1$$

lo que conduce a la misma conclusión anterior. Es obvio que esto se esperaba, puesto que la corrección de Yates siempre *reduce* el valor de χ^2. Aquí, el valor P es $P(\chi^2 \geq 2.1) = 0.15$.

Debemos observar que si se usa la aproximación χ^2 a pesar de que las frecuencias sean demasiado pequeñas, se obtendría

$$\chi^2 = \frac{(2 - 3)^2}{3} + \frac{(5 - 3)^2}{3} + \frac{(4 - 3)^2}{3} + \frac{(1 - 3)^2}{3} = 3.33$$

con un valor P de 0.34.

Debido a que para $\nu = 4 - 1 = 3$, $\chi^2_{0.95} = 7.81$, se llegaría a las mismas conclusiones que antes. Desafortunadamente, la aproximación χ^2 para frecuencias pequeñas es pobre; en consecuencia, cuando no sea recomendable combinar frecuencias se necesita acudir a los métodos exactos de probabilidad que implican la distribución multinomial.

7.41. En 360 lanzamientos de un par de dados, se observan 74 "sietes" y 24 "onces". Con base en un nivel de significancia de 0.05, probar la hipótesis de que los dados no están cargados y determine el valor P de los resultados observados.

Un par de dados puede caer de 36 maneras. Un siete puede ocurrir de 6 maneras, un once de 2.

$= \frac{6}{36} = \frac{1}{6}$ y $P(\text{once}) = \frac{2}{36} = \frac{1}{18}$. Tenemos que, en 360 lanzamientos se deberían esperar $\frac{1}{6}(360) = 60$ sietes y $\frac{1}{18}(360) = 20$ onces, por lo que

$$\chi^2 = \frac{(74 - 60)^2}{60} + \frac{(24 - 20)^2}{20} = 4.07$$

con un valor P de 0.044.

Para $\nu = 2 - 1 = 1$, $\chi^2_{0.95} = 3.84$. Por consiguiente, en vista de que $4.07 > 3.84$, se debería estar inclinado a rechazar la hipótesis de que los dados no están cargados. Sin embargo, si se emplea la corrección de Yates, se encuentra que

$$\chi^2\,(\text{corregida}) = \frac{(|74 - 60| - 0.5)^2}{60} + \frac{(|24 - 20| - 0.5)^2}{20} = \frac{(13.5)^2}{60} + \frac{(3.5)^2}{20} = 3.65$$

Con un valor P de 0.056.

En consecuencia, con base en la χ^2 corregida, no es posible rechazar la hipótesis a un nivel de 0.05.

Por lo general, para muestras grandes como las de este ejemplo, los resultados que se logran mediante la corrección de Yates son más confiables que los resultados no corregidos. Sin embargo, puesto que incluso el valor corregido de χ^2 se ubica tan cerca del valor crítico, se vacila en tomar una u otra decisión. En tales casos, quizá sea mejor aumentar el tamaño de la muestra y tomar más observaciones si interesa en especial el nivel de 0.05. En caso contrario, es posible rechazar la hipótesis a otro nivel (como 0.10).

7.42. Una encuesta de 320 familias con 5 niños cada una reveló la distribución de varones y mujeres que se muestra en la tabla 7-12. *a*) ¿Es el resultado consistente con la hipótesis de que los nacimientos de varones y de mujeres son igualmente probables? *b*) ¿Cuál es el valor P de los resultados del muestreo?

a)

Tabla 7-12

Número de niños y niñas	5 niños 0 niñas	4 niños 1 niña	3 niños 2 niñas	2 niños 3 niñas	1 niño 4 niñas	0 niños 5 niñas	TOTAL
Número de familias	18	56	110	88	40	8	320

Sea p = probabilidad de un nacimiento masculino y $q = 1 - p$ = probabilidad de un nacimiento femenino. En consecuencia, las probabilidades de (5 niños), (4 niños y 1 niña), . . . , (5 niñas) están dadas por los términos de la expansión binomial

$$(p + q)^5 = p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5$$

Si $p = q = \frac{1}{2}$, se tiene

$$P(\text{5 niños y 0 niñas}) = \left(\frac{1}{2}\right)^5 = \frac{1}{32} \qquad P(\text{2 niños y 3 niñas}) = 10\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 = \frac{10}{32}$$

$$P(\text{4 niños y 1 niña}) = 5\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right) = \frac{5}{32} \qquad P(\text{1 niño y 4 niñas}) = 5\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^4 = \frac{5}{32}$$

$$P(\text{3 niños y 2 niñas}) = 10\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{10}{32} \qquad P(\text{0 niños y 5 niñas}) = \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

Por tanto, se obtiene el número esperado de familias con 5, 4, 3, 2, 1 y 0 niños, respectivamente, al multiplicar las probabilidades anteriores por 320, los resultados son 10, 50, 100, 100, 50, 10. Por tanto,

$$\chi^2 = \frac{(18 - 10)^2}{10} + \frac{(56 - 50)^2}{50} + \frac{(100 - 100)^2}{100} + \frac{(88 - 100)^2}{100} + \frac{(40 - 50)^2}{50} + \frac{(8 - 10)^2}{10} = 12.0$$

Puesto que $\chi^2_{0.95} = 11.1$ y $\chi^2_{0.99} = 15.1$ para $\nu = 6 - 1 = 5$ grados de libertad, es posible rechazar la hipótesis a un nivel de significancia de 0.05, pero no a uno de 0.01. En consecuencia, concluimos que los resultados son probablemente significativos, y que los nacimientos masculinos y femeninos no tienen igual probabilidad de ocurrir.

b) El valor P es $P(\chi^2 \geq 12.0) = 0.035$.

BONDAD DE AJUSTE

7.43. Utilizar la prueba ji-cuadrada para determinar la bondad de ajuste de los datos del problema 7.30.

$$\chi^2 = \frac{(38 - 33.2)^2}{33.2} + \frac{(144 - 161.9)^2}{161.9} + \frac{(342 - 316.2)^2}{316.2} + \frac{(287 - 308.7)^2}{308.7} + \frac{(164 - 150.7)^2}{150.7} + \frac{(25 - 29.4)^2}{29.4}$$

$$= 7.45$$

Puesto que el número de parámetros que se usa para estimar las frecuencias esperadas es $m = 1$ (de manera específica, el parámetro p de la distribución binomial), $\nu = k - 1 - m = 6 - 1 - 1 = 4$.

Para $\nu = 4$, $\chi^2_{0.95} = 9.49$. Por tanto, se deduce que el ajuste de los datos es muy bueno.

Para $\nu = 4$, $\chi^2_{0.05} = 0.711$. En consecuencia, puesto que $\chi^2 = 7.54 > 0.711$, el ajuste no es tan preciso como para ser increíble.

El valor P es $P(\chi^2 \geq 7.45) = 0.11$.

7.44. Determinar la bondad de ajuste de los datos del problema 7.32.

$$\chi^2 = \frac{(5 - 4.13)^2}{4.13} + \frac{(18 - 20.68)^2}{20.68} + \frac{(42 - 38.92)^2}{38.92} + \frac{(27 - 27.71)^2}{27.71} + \frac{(8 - 7.43)^2}{7.43} = 0.959$$

Puesto que el número de parámetros que se utiliza para estimar las frecuencias esperadas es $m = 2$ (esto es, la media μ y la desviación estándar σ de la distribución normal), $\nu = k - 1 - m = 5 - 1 - 2 = 2$.

Para $\nu = 2$, $\chi^2_{0.95} = 5.99$. Por tanto, se concluye que el ajuste de los datos es muy preciso.

Para $\nu = 2$, $\chi^2_{0.05} = 0.103$. En consecuencia, $\chi^2 = 0.959 > 0.103$ se concluye que el ajuste no es "muy bueno".

El valor P es $P(\chi^2 \geq 0.959) = 0.62$.

TABLAS DE CONTINGENCIA

7.45. Resolver el problema 7.13 con base en el empleo de la prueba ji-cuadrada.

En la tabla 7-13 se presentan las condiciones del problema. Con base en la hipótesis nula H_0 de que el suero no tiene efecto, se debería esperar que 70 personas en cada uno de los grupos se recuperan y 30 en cada grupo no se recuperan, como se indica en la tabla 7-14. Observemos que H_0 es equivalente a la afirmación de que la recuperación es independiente del uso del suero, esto es, las clasificaciones son independientes.

Tabla 7-13
Frecuencias observadas

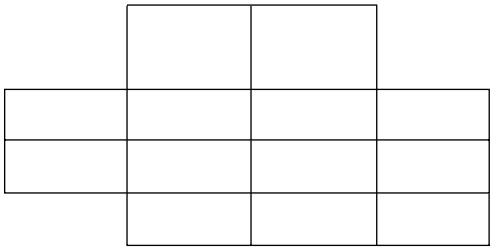
	Se recuperan	No se recuperan	TOTAL
Grupo *A* (utilizando suero)	75	25	100
Grupo *B* (no utilizando suero)	65	35	100
TOTAL	140	60	200

Tabla 7-14
Frecuencias esperadas bajo H_0

	Se recuperan	No se recuperan	TOTAL
Grupo *A* (utilizando suero)	70	30	100
Grupo *B* (no utilizando suero)	70	30	100
TOTAL	140	60	200

$$\chi^2 = \frac{(75 - 70)^2}{70} + \frac{(65 - 70)^2}{70} + \frac{(25 - 30)^2}{30} + \frac{(35 - 30)^2}{30} = 2.38$$

Para determinar el número de grados de libertad, veamos la tabla 7-15, que es la misma que las tablas 7-13 y 7-14, excepto que sólo se muestran los totales. Es claro que se tiene la libertad de poner sólo un número en cual-



$$\chi^2(\text{corregida}) = \frac{(|75 - 70| - 0.5)^2}{70} + \frac{(|65 - 70| - 0.5)^2}{70} + \frac{(|25 - 30| - 0.5)^2}{30} + \frac{(|35 - 30| - 0.5)^2}{30} = 1.93$$

con un valor P de 0.16.

En consecuencia, las conclusiones dadas en el problema 7.45 son válidas. Esto debió haberlo notado de inmediato al observar que la corrección de Yates siempre decrece el valor χ^2 y aumenta el valor P.

7.47. La tabla 7-16 muestra los números de estudiantes que fueron aprobados y reprobados por tres instructores: Sr. X, Sr. Y y Sr. Z. Probar la hipótesis de que las proporciones de estudiantes que los tres instructores no aprobaron son iguales.

Tabla 7-16

Frecuencias observadas

	Sr. X	Sr. Y	Sr. Z	TOTAL
Aprobados	50	47	56	153
Reprobados	5	14	8	27
TOTAL	55	61	64	180

Con base en la hipótesis H_0 de que las proporciones de estudiantes no aprobados por los tres instructores son iguales, hubieran reprobado $27/180 = 15\%$ de los estudiantes y aprobado 85% de ellos. Las frecuencias esperadas con base en H_0 se muestran en la tabla 7-17.

$$\chi^2 = \frac{(50 - 46.75)^2}{46.75} + \frac{(47 - 51.85)^2}{51.85} + \frac{(56 - 54.40)^2}{54.40} + \frac{(5 - 8.25)^2}{8.25} + \frac{(14 - 9.15)^2}{9.15} + \frac{(8 - 9.60)^2}{9.60} = 4.84$$

Tabla 7-17

Frecuencias esperadas bajo H_0

	Sr. X	Sr. Y	Sr. Z	TOTAL
Aprobados	85% de 55 = 46.75	85% de 61 = 51.85	85% de 64 = 54.40	153
Reprobados	15% de 55 = 8.25	15% de 61 = 9.15	15% de 64 = 9.60	27
TOTAL	55	61	64	180

Para determinar el número de grados de libertad, veamos la tabla 7-18, que es igual a las tablas 7-16 y 7-17, excepto que sólo se muestran los totales. Es claro que se tiene la libertad de poner sólo un número en una celda vacía de la primera columna y un número en una celda vacía en la segunda o la tercera columna, después de lo cual los números de las celdas restantes serán determinados de manera única a partir de los totales indicados. Por tanto, en este caso hay 2 grados de libertad.

Tabla 7-18

	Sr. X	Sr. Y	Sr. Z	TOTAL
Aprobados				153
Reprobados				27
TOTAL	55	61	64	180

Como $\chi^2_{0.95} = 5.99$, no es posible rechazar H_0 a un nivel de 0.05. No obstante, observemos que puesto que $\chi^2_{0.90} = 4.61$, se puede rechazar H_0 a un nivel de 0.10 si se está dispuesto a tomar el riesgo de 1 en 10 posibilidades de equivocarse. El valor P de las frecuencias observadas es $P(\chi^2 \geq 4.84) = 0.089$.

7.48. Demostrar que en una tabla de contingencia de $h \times k$ ($h > 1$, $k > 1$), el número de grados de libertad está dado por $(h - 1)(k - 1)$.

Hay $h + k - 1$ totales independientes de las entradas hk. De ello se infiere que el número de grados de libertad es

$$hk - (h + k - 1) = (h - 1)(k - 1)$$

según se requiere. El resultado se mantiene si se conocen los parámetros poblacionales necesarios para obtener frecuencias teóricas; en el caso contrario se necesita ajustar como se describe en *b*), página 221.

7.49. La tabla 7-19 representa una tabla de contingencia de 2×2 general. Demostrar que

$$\chi^2 = \frac{n(a_1b_2 - a_2b_1)^2}{n_1n_2n_An_B}$$

Tabla 7-19

Resultados observados

	I	II	TOTAL
A	a_1	a_2	n_A
B	b_1	b_2	n_B
TOTAL	n_1	n_2	n

Tabla 7-20

Resultados esperados

	/	/	n_A
	/	/	

Al igual que en el problema 7.45, los resultados esperados con base en una hipótesis nula se muestran en la tabla 7-20. En consecuencia,

$$\chi^2 = \frac{(a_1 - n_1n_A/n)^2}{n_1n_A/n} + \frac{(a_2 - n_2n_A/n)^2}{n_2n_A/n} + \frac{(b_1 - n_1n_B/n)^2}{n_1n_B/n} + \frac{(b_2 - n_2n_B/n)^2}{n_2n_B/n}$$

Pero $$a_1 - \frac{n_1n_A}{n} = a_1 - \frac{(a_1 + b_1)(a_1 + a_2)}{a_1 + b_1 + a_2 + b_2} = \frac{a_1b_2 - a_2b_1}{n}$$

De manera similar, $$a_2 - \frac{n_2n_A}{n} = b_1 - \frac{n_1n_B}{n} = b_2 - \frac{n_2n_B}{n} = \frac{a_1b_2 - a_2b_1}{n}$$

En consecuencia, se escribe

$$\chi^2 = \frac{n}{n_1n_A}\left(\frac{a_1b_2 - a_2b_1}{n}\right)^2 + \frac{n}{n_2n_A}\left(\frac{a_1b_2 - a_2b_1}{n}\right)^2 + \frac{n}{n_1n_B}\left(\frac{a_1b_2 - a_2b_1}{n}\right)^2 + \frac{n}{n_2n_B}\left(\frac{a_1b_2 - a_2b_1}{n}\right)^2$$

lo que se simplifica en

1) $$\chi^2 = \frac{n(a_1b_2 - a_2b_1)^2}{n_1n_2n_An_B} = \frac{n\Delta^2}{n_1n_2n_An_B}$$

donde $\Delta = a_1b_2 - a_2b_1$, $n = a_1 + a_2 + b_1 + b_2$, $n_1 = a_1 + b_1$, $n_2 = a_2 + b_2$, $n_A = a_1 + a_2$, $n_B = b_1 + b_2$. Si se aplica la corrección de Yates, se sustituye 1) por

2) $$\chi^2(\text{corregida}) = \frac{n\left(|\Delta| - \frac{1}{2}n\right)^2}{n_1n_2n_An_B}$$

7.50. Ilustrar el resultado del problema 7.49 con los datos del problema 7.45.

En el problema 7.45, $a_1 = 75$, $a_2 = 25$, $b_1 = 65$, $b_2 = 35$, $n_1 = 140$, $n_2 = 60$, $n_A = 100$, $n_B = 100$, y $n = 200$; entonces 1) del problema 7.49 obtenemos

$$\chi^2 = \frac{200[(75)(35) - (25)(65)]^2}{(140)(60)(100)(100)} = 2.38$$

Con la corrección de Yates, el resultado es igual que el del problema 7.46:

$$\chi^2(\text{corregida}) = \frac{n\left(|a_1b_2 - a_2b_1| - \frac{1}{2}n\right)^2}{n_1n_2n_An_B} = \frac{200[|(75)(35) - (25)(65)| - 100]^2}{(140)(60)(100)(100)} = 1.93$$

7.51. Demostrar, empleando la aproximación normal, que una prueba ji-cuadrada que relaciona dos proporciones muestrales equivale a una prueba de significancia de diferencias en proporciones (vea la página 217).

Sean P_1 y P_2 las dos proporciones muestrales y p la proporción poblacional. Con referencia al problema 7.49, se tiene que

1) $$P_1 = \frac{a_1}{n_1}, \qquad P_2 = \frac{a_2}{n_2}, \qquad 1 - P_1 = \frac{b_1}{n_1}, \qquad 1 - P_2 = \frac{b_2}{n_2}$$

2) $$p = \frac{n_A}{n}, \qquad 1 - p = q = \frac{n_B}{n}$$

$$a_1 = n_1P_1, \qquad a_2 = n_2P_2, \qquad b_1 = n_1(1 - P_1), \qquad = n_2(1 - P_2)$$

4) $$n_A = np, \qquad n_B = nq$$

Con base en 3) y 4), y del problema 7.49, se tiene que

$$\chi^2 = \frac{n(a_1b_2 - a_2b_1)^2}{n_1n_2n_An_B} = \frac{n[n_1P_1n_2(1 - P_2) - n_2P_2n_1(1 - P_1)]^2}{n_1n_2npnq}$$

$$= \frac{n_1n_2(P_1 - P_2)^2}{npq} = \frac{(P_1 - P_2)^2}{pq(1/n_1 + 1/n_2)} \qquad (\text{donde } n = n_1 + n_2)$$

que es el cuadrado del estadístico Z dado en (10) en la página 217.

COEFICIENTE DE CONTINGENCIA

7.52. Determinar el coeficiente de contingencia con los datos de la tabla de contingencia del problema 7.45.

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{2.38}{2.38 + 200}} = \sqrt{0.01176} = 0.1084$$

7.53. Calcular el valor máximo de C de todas las tablas de 2×2 que podrían surgir en el problema 7.13.

El valor máximo de C ocurre cuando las dos clasificaciones son perfectamente dependientes o asociadas. En estos casos, todos aquellos que toman el suero se recuperarán y los que no lo toman no se recuperan. La tabla de contingencia es la 7-21.

Tabla 7-21

	Se recuperan	No se recuperan	TOTAL
Grupo *A* (se utilizó suero)	100	0	100
Grupo *B* (no se utilizó suero)	0	100	100
TOTAL	100	100	200

En razón de que las frecuencias esperadas de las celdas, bajo el supuesto de independencia total, son todas iguales a 50,

$$\chi^2 = \frac{(100 - 50)^2}{50} + \frac{(0 - 50)^2}{50} + \frac{(0 - 50)^2}{50} + \frac{(100 - 50)^2}{50} = 200$$

En consecuencia, el valor máximo de C es $\sqrt{\chi^2/(\chi^2 + n)} = \sqrt{200/(200 + 200)} = 0.7071$.

En general, para dependencia perfecta en una tabla de contingencia donde los números de filas y columnas son ambos igual a k, la única frecuencia de celda no cero ocurre en la diagonal desde el extremo izquierdo superior al derecho inferior. En estos casos, $C_{\text{máx}} = \sqrt{(k - 1)/k}$.

PROBLEMAS DIVERSOS

7.54. Un instructor aplica una prueba breve que implica 10 preguntas falso-verdadero. Para probar la hipótesis de que el estudiante adivina, se adopta la siguiente regla de decisión: *a*) si 7 o más son correctos, el estudiante no ha adivinado; *b*) si menos de 7 son correctos, el estudiante sí ha adivinado. Calcular la probabilidad de rechazar la hipótesis cuando ésta es correcta.

Sea p = probabilidad de que se contesta correctamente una pregunta.

La probabilidad de responder a x preguntas de 10 es ${}_{10}C_x p^x q^{10-x}$, donde $q = 1 - p$.

En consecuencia, según la hipótesis $p = 0.5$ (esto es, él está adivinando),

$$= P(\text{7 correctas}) + P(\text{8 correctas}) + P(\text{9 correctas}) + P(\text{10 correctas})$$

$$= {}_{10}C_7\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^3 + {}_{10}C_8\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^2 + {}_{10}C_9\left(\frac{1}{2}\right)^9\left(\frac{1}{2}\right) + {}_{10}C_{10}\left(\frac{1}{2}\right)^{10} = 0.1719$$

Por tanto, la probabilidad de concluir que el estudiante no está adivinando cuando de hecho sí lo hace es de 0.1719. Observemos que ésta es la probabilidad de un error del tipo I.

7.55. En el problema 7.54, Calcular la probabilidad de aceptar la hipótesis $p = 0.5$ cuando en realidad $p = 0.7$.

Con base en la hipótesis $p = 0.7$,

$P(\text{menos de 7 correctas}) = 1 - P(\text{7 o más correctas})$

$$= 1 - [{}_{10}C_7(0.7)^7(0.3)^3 + {}_{10}C_8(0.7)^8(0.3)^2 + {}_{10}C_9(0.7)^9(0.3) + {}_{10}C_{10}(0.7)^{10}] = 0.3504$$

7.56. En el problema 7.54, determinar la probabilidad de aceptar la hipótesis $p = 0.5$ cuando en realidad *a*) $p = 0.6$, *b*) $p = 0.8$, *c*) $p = 0.9$, *d*) $p = 0.4$, *e*) $p = 0.3$, *f*) $p = 0.2$, *g*) $p = 0.1$.

a) Si $p = 0.6$, la probabilidad que se requiere está dada por

$$1 - [P(\text{7 correctas}) + P(\text{8 correctas}) + P(\text{9 correctas}) + P(\text{10 correctas})]$$

$$= 1 - [{}_{10}C_7(0.6)^7(0.4)^3 + {}_{10}C_8(0.6)^8(0.4)^2 + {}_{10}C_9(0.6)^9(0.4) + {}_{10}C_{10}(0.6)^{10}] = 0.618$$

Es posible encontrar de manera similar los resultados para *b*), *c*), . . . , *g*) y se indican en la tabla 7-22 junto con el valor correspondiente a $p = 0.7$ que se encontró en el problema 7.55. Observe que se denota la probabilidad por β (probabilidad de un error del tipo II). También se incluye la entrada para $p = 0.5$, dada por $\beta = 1 - 0.1719 = 0.828$ del problema 7.54.

Tabla 7-22

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
β	1.000	0.999	0.989	0.945	0.828	0.618	0.350	0.121	0.013

7.57. Utilizar el problema 7.56 para construir la gráfica de β *vs*. p, la curva característica de operación de la regla de decisión en el problema 7.54.

La gráfica que se requiere se muestra en la figura 7-14. Observemos la similitud con la curva CO del problema 7.27.

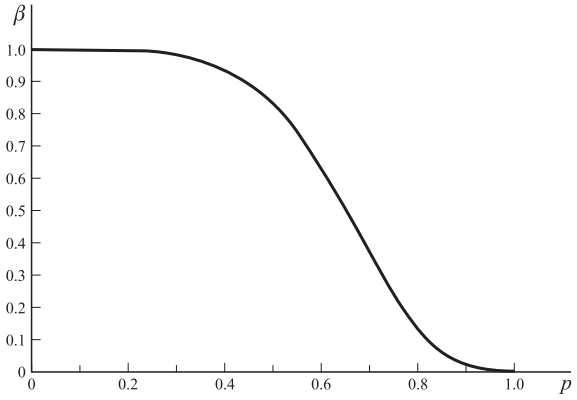


Figura 7-14

Si se hubiera graficado $(1 - \beta)$ *vs*. p, se hubiera obtenido la *curva de potencia* de la regla de decisión.
La gráfica indica que la regla de decisión dada es *poderosa* para rechazar $p = 0.5$ cuando en realidad $p \geq 0.8$.

7.58. Una moneda que se lanza 6 veces, en todas ellas resulta cara. ¿Es posible concluir que la moneda está cargada a un nivel de significancia de *a*) 0.05, *b*) 0.01? Considerar tanto una prueba de una cola como una de dos colas.

$$= P(x \text{ caras en 6 lanzamientos}) = {}_6C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{6-x} = \frac{{}_6C_x}{64}$$

En consecuencia, las probabilidades de 0, 1, 2, 3, 4, 5 y 6 caras están dadas, respectivamente, por $\frac{1}{64}, \frac{6}{64}, \frac{15}{64}, \frac{20}{64}, \frac{15}{64}, \frac{6}{64}$ y $\frac{1}{64}$.

Prueba de una cola
En este caso se debe decidir entre la hipótesis (H_0: $p = 0.5$) y (H_1: $p > 0.5$). Puesto que $P(6 \text{ caras}) = \frac{1}{64} = 0.01562$ y $P(5 \text{ o } 6 \text{ caras}) = \frac{6}{64} + \frac{1}{64} = 0.1094$, es posible rechazar H_0 a un nivel de 0.05 pero no a uno de 0.01 (esto es, el resultado observado es significativo a un nivel de 0.05 pero no a uno de 0.01).

Prueba de dos colas
Aquí se desea decidir entre la hipótesis (H_0: $p = 0.5$) y (H_1: $p \neq 0.5$). Puesto que $P(0 \text{ o } 6 \text{ caras}) = \frac{1}{64} + \frac{1}{64} = 0.03125$, podemos rechazar H_0 a un nivel de 0.05 pero no a uno de 0.01.

7.59. Resolver el problema 7.58 si la moneda resulta en caras 5 veces.

Prueba de una cola
En razón de que $P(5 \text{ o } 6 \text{ caras}) = \frac{6}{64} + \frac{1}{64} = \frac{7}{64} = 0.1094$, no se puede rechazar H_0 a un nivel de 0.05 o 0.01.

Prueba de dos colas
Puesto que $P(0 \text{ o } 1 \text{ o } 5 \text{ o } 6 \text{ caras}) = 2\left(\frac{7}{64}\right) = 0.2188$, no se puede rechazar H_0 a un nivel de 0.05 o 0.01.

7.60. Demostrar que la prueba ji-cuadrada que implica sólo dos categorías equivale a la prueba de significancia para proporciones (página 216).

Si P es la proporción muestral de la categoría I, p es la proporción poblacional y n es la frecuencia total, se puede describir la situación por medio de la tabla 7-23. Entonces, por definición,

$$\chi^2 = \frac{(nP - np)^2}{np} + \frac{[n(1 - P) - n(1 - p)]^2}{nq}$$

$$= \frac{n^2(P - p)^2}{np} + \frac{n^2(P - p)^2}{nq} = n(P - p)^2\left(\frac{1}{p} + \frac{1}{q}\right) = \frac{n(P - p)^2}{pq} = \frac{(P - p)^2}{pq/n}$$

que es el cuadrado del estadístico Z (5) de la página 216.

Tabla 7-23

	I	II	TOTAL
Frecuencia observada	nP	$n(1 - P)$	n
Frecuencia esperada	np	$n(1 - p) = nq$	n

7.61. Se supone que $X_1, X_2, \ldots, X_k$ tiene una distribución multinomial, con frecuencias respectivas esperadas np_1, $np_2, \ldots, np_k$,. Sean $Y_1, Y_2, \ldots, Y_k$ variables mutuamente independientes de la distribución de Poisson, con parámetros respectivos $\lambda_1 = np_1, \lambda_2 = np_2, \ldots, \lambda_k = np_k$,. Demostrar que la distribución condicional de las Y si

$$Y_1 + Y_2 + \cdots + Y_k = n$$

es precisamente la distribución multinomial de las X.

Para la función de probabilidad conjunta de las Y, tenemos

$$1) \qquad P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_k = y_k) = \left[\frac{(np_1)^{y_1}e^{-np_1}}{y_1!}\right]\left[\frac{(np_2)^{y_2}e^{-np_2}}{y_2!}\right]\cdots\left[\frac{(np_k)^{y_k}e^{-np_k}}{y_k!}\right]$$

$$= \frac{n^{y_1+y_2+\cdots+y_k}p_1^{y_1}p_2^{y_2}\cdots p_k^{y_k}}{y_1!y_2!\cdots y_k!}e^{-n}$$

$+ p_2 + \cdots + p_k = 1$. La distribución condicional que se busca está dada por

2) $$P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_k = y_k | Y_1 + Y_2 + \cdots + Y_k = n) = \frac{P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_k = y_k \text{ y } Y_1 + Y_2 + \cdots + Y_k = n)}{P(Y_1 + Y_2 + \cdots + Y_k = n)}$$

Ahora, a partir de 1), el numerador en 2) tiene el valor

$$\frac{n^n p_1^{y_1} p_2^{y_2} \cdots p_k^{y_k}}{y_1! y_2! \cdots y_k!} e^{-n}$$

En cuanto al denominador, se sabe del problema 4.94, página 146, que $Y_1 + Y_2 + \cdots + Y_k$ es en sí una variable de Poisson con parámetro $np_1 + np_2 + \cdots + np_k = n$. Por tanto, el denominador tiene el valor

$$\frac{n^n e^{-n}}{n!}$$

En consecuencia, 2) se convierte en

$$P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_k = y_k | Y_1 + Y_2 + \cdots + Y_k = n) = \frac{n!}{y_1! y_2! \cdots y_k!} p_1^{y_1} p_2^{y_2} \cdots p_k^{y_k}$$

que es justo la distribución multinomial de las X [comparar con (16), página 112].

7.62. Utilizar el resultado del problema 7.61 para demostrar que χ^2, como se define en (21), página 220, es aproximadamente una distribución ji-cuadrada.

En su forma actual (21) es difícil de tratar porque las X distribuidas de manera multinomial son dependientes, dada la restricción (22). Sin embargo, el problema 7.61 muestra que es posible sustituir las X por las Y *independientes* con distribución de Poisson si está dado que $Y_1 + Y_2 + \cdots + Y_k = n$. Por tanto, (21) se reescribe como

1) $$\chi^2 = \left(\frac{Y_1 - \lambda_1}{\sqrt{\lambda_1}}\right)^2 + \left(\frac{Y_2 - \lambda_2}{\sqrt{\lambda_2}}\right)^2 + \cdots + \left(\frac{Y_k - \lambda_k}{\sqrt{\lambda_k}}\right)^2$$

Cuando $n \to \infty$, todas las λ tienden a ∞, y el teorema del límite central para la distribución de Poisson [(14), página 112] da como resultado

2) $$\chi^2 \approx Z_1^2 + Z_2^2 + \cdots + Z_k^2$$

donde las Z son variables normales *independientes* con media 0 y varianza 1 cuya distribución es condicional dependiendo del evento

3) $$\sqrt{\lambda_1} Z_1 + \sqrt{\lambda_2} Z_2 + \cdots + \sqrt{\lambda_k} Z_k = 0 \quad \text{o bien} \quad \sqrt{p_1} Z_1 + \sqrt{p_2} Z_2 + \cdots + \sqrt{p_k} Z_k = 0$$

o, puesto que las variables aleatorias son continuas,

4) $$|\sqrt{p_1} Z_1 + \sqrt{p_2} Z_2 + \cdots + \sqrt{p_k} Z_k| < \epsilon$$

Sea $F_\nu(x)$ la función de distribución acumulativa de una variable ji-cuadrada con ν grados de libertad. Entonces, lo que queremos demostrar es:

5) $$P\left(Z_1^2 + Z_2^2 + \cdots + Z_k^2 \le x \,|\, |\sqrt{p_1} Z_1 + \sqrt{p_2} Z_2 + \cdots + \sqrt{p_k} Z_k| < \epsilon\right) = \frac{P\left(Z_1^2 + Z_2^2 + \cdots + Z_k^2 \le x \text{ y } |\sqrt{p_1} Z_1 + \sqrt{p_2} Z_2 + \cdots + \sqrt{p_k} Z_k| < \epsilon\right)}{P(|\sqrt{p_1} Z_1 + \sqrt{p_2} Z_2 + \cdots + \sqrt{p_k} Z_k| < \epsilon)} = F_\nu(x)$$

para un valor adecuado de ν.

Es fácil establecer 5) si se usa intuición geométrica. En primer lugar, el teorema 4-3 demuestra que la distribución no condicional de $Z_1^2 + Z_2^2 + \cdots + Z_k^2$ es ji-cuadrada con k grados de libertad. En consecuencia, puesto que la función de densidad de cada Z_j es $(2\pi)^{-1/2} e^{-z^2/2}$,

$$F_k(x) = (2\pi)^{-k/2} \int \cdots \int_{+z_2^2+\cdots+z_k^2 \le x} e^{-(z_1^2+z_2^2+\cdots+z_k^2)/2} dz_1 dz_2 \cdots dz_k$$

Además, se tiene del numerador en 5):

$$\text{7)} \qquad \text{Numerador} = (2\pi)^{-k/2} \int \cdots \int_{\substack{z_1^2+z_2^2+\cdots+z_k^2 \le x, \\ |\sqrt{p_1}z_1+\sqrt{p_2}z_2+\cdots+\sqrt{p_k}z_k|<\epsilon}} e^{-(z_1^2+z_2^2+\cdots+z_k^2)/2} dz_1 dz_2 \cdots dz_k$$

Recordemos de geometría analítica que en el espacio tridimensional, $x_1^2 + x_2^2 + x_3^2 \le a^2$ representa un sólido esférico de radio a centrado en el origen, mientras que $\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = 0$ es un plano que pasa por el origen y cuya normal es el vector unitario $(\alpha_1, \alpha_2, \alpha_3)$. La figura 7-15 muestra la intersección de los dos cuerpos. Es obvio que cuando una función que depende sólo de la distancia desde el origen, esto es,

$$f(r) \qquad \text{donde} \qquad r = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

se integra sobre el área circular, o a lo largo de una rebanada delgada que yace sobre esa área, el valor de la integral es totalmente independiente de los cosenos direccionales α_1, α_2, α_3. Es decir, todos los planos cortantes que pasan por el origen producen la misma integral.

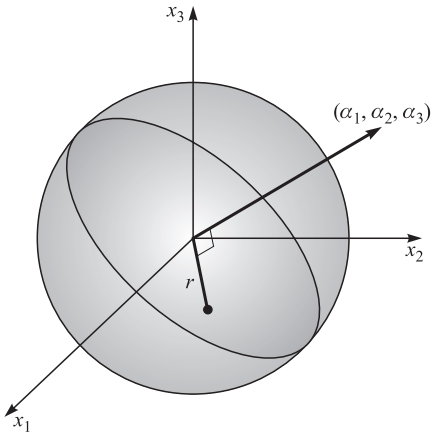


Figura 7-15

Por analogía, se concluye que en 7), donde $e^{-r^2/2}$ se integra sobre la intersección de una *hiperesfera* alrededor del origen y un *hiperplano* que pasa por el origen, es posible dar a las p cualquier valor conveniente. Se elige

$$p_1 = p_2 = \cdots = p_{k-1} = 0, \qquad p_k = 1$$

y se obtiene

$$\text{8)} \qquad \text{Numerador} = (2\pi)^{-k/2} \int \cdots \int_{z_1^2+z_2^2+\cdots+z_{k-1}^2 \le x} e^{-(z_1^2+z_2^2+\cdots+z_{k-1}^2)/2} dz_1 dz_2 \cdots dz_{k-1}(2\epsilon)$$

$$= (2\pi)^{-1/2} F_{k-1}(x)(2\epsilon)$$

mediante el uso de 6). El factor 2ϵ es el grosor de la rebanada.

Para evaluar el denominador en 5), se nota que la variable aleatoria

$$W = \sqrt{p_1}Z_1 + \sqrt{p_2}Z_2 + \cdots + \sqrt{p_k}Z_k$$

es normal (porque es una combinación lineal de las Z independientes y normales), y que

$$E(W) = \sqrt{p_1}(0) + \sqrt{p_2}(0) + \cdots + \sqrt{p_k}(0) = 0$$
$$\text{Var}(W) = p_1(1) + p_2(1) + \cdots + p_k(1) = 1$$

Por tanto, la función de densidad de W es $\phi(w) = (2\pi)^{-1/2} e^{-w^2/2}$, y

$$\text{9)} \qquad \text{Denominador} = P(|W| < \epsilon) = \phi(0)(2\epsilon) = (2\pi)^{-1/2}(2\epsilon)$$

Cuando se divide 8) entre 9) se obtiene el resultado deseado, donde $\nu = k - 1$.

La "demostración" anterior (que se puede hacer rigurosa) muestra de manera incidental que cada restricción lineal impuesta en las Z, y, por tanto, en las Y o X, reduce el número de grados de libertad en χ^2 por 1. Esto proporciona la base de las reglas dadas en la página 221.

PROBLEMAS SUPLEMENTARIOS

PRUEBAS DE MEDIAS Y PROPORCIONES CON EL EMPLEO DE DISTRIBUCIONES NORMALES

7.63. Una urna contiene canicas rojas y azules. Para probar la hipótesis de proporciones iguales de estos colores, se acuerda muestrear 64 canicas con reemplazo, anotar los colores extraídos y adoptar la siguiente regla de decisión: 1) aceptar la hipótesis si se extraen entre 28 y 36 canicas rojas; 2) rechazarla en otro caso. *a*) Determine la probabilidad de rechazar la hipótesis cuando en realidad es correcta. *b*) Interprete gráficamente la regla de decisión y el resultado que obtuvo en *a*).

7.64. *a*) ¿Cuál regla de decisión se debería adoptar en el problema 7.63 si se requiere que la probabilidad de rechazar la hipótesis cuando ésta es correcta hasta 0.01 como máximo? En otras palabras, se quiere un nivel de significancia de 0.01. *b*) ¿A cuál nivel de confianza se debería aceptar la hipótesis? *c*) ¿Cuál sería la regla de decisión si se aplica un nivel de significancia de 0.05?

7.65. Suponga que en el problema 7.63 se desea probar la hipótesis de que hay una proporción mayor de canicas rojas que azules. *a*) ¿Cuál sería la hipótesis nula y cuál la hipótesis alternativa? *b*) ¿Se debe emplear una prueba de una cola o de dos? ¿Por qué? *c*) ¿Cuál regla de decisión se debe adoptar si el nivel de significancia es de 0.05? *d*) ¿Cuál debe ser la regla de decisión si el nivel de significancia es de 0.01?

7.66. Un par de dados se lanza 100 veces, y se observa que la suma siete aparece 23 veces. Pruebe la hipótesis de que los dados no están cargados con base en *a*) una prueba de dos colas y *b*) una prueba de una cola, las dos con un nivel de significancia de 0.05. Explique las razones, si las hay, de preferir una de estas pruebas más que la otra.

7.67. Resuelva el problema 7.66 si el nivel de significancia es 0.01.

7.68. Un fabricante afirma que por lo menos 95% del equipo que entregó a una fábrica cumple con las especificaciones. Una revisión de una muestra de 200 piezas de equipo reveló que 18 presentaban defectos. Pruebe la afirmación del fabricante a un nivel de significancia de *a*) 0.01, *b*) 0.05.

7.69. Se ha encontrado por experiencia que la media de una fuerza de rompimiento de una marca particular de hilo es de 9.72 onzas con una desviación estándar de 1.4 onzas. En fechas recientes, una muestra de 36 piezas de hilo mostró que la media de una fuerza de rompimiento era de 8.93 onzas. ¿Es posible concluir que el hilo se ha vuelto inferior a un nivel de significancia de *a*) 0.05, *b*) 0.01?

7.70. En un examen que se aplicó a estudiantes en un gran número de escuelas diferentes, la calificación media fue 74.5 y la desviación estándar 8.0. En una escuela determinada, donde 200 estudiantes hicieron el examen, la calificación media fue 75.9. Analice la significancia de este resultado a un nivel de 0.05 desde la perspectiva de *a*) una prueba de una cola, *b*) una prueba de dos colas. Explique con detalle las conclusiones con base en estas pruebas.

7.71. Resuelva el problema 7.70 si el nivel de significancia es de 0.01.

PRUEBAS EN LAS QUE INTERVIENEN DIFERENCIAS ENTRE MEDIAS Y ENTRE PROPORCIONES

7.72. Una muestra de 100 focos eléctricos que produce el fabricante *A* mostró que la media del tiempo de vida es de 1 190 horas y una desviación estándar de 90 horas. Una muestra de 75 focos que fabrica la empresa *B* mostró que la media del tiempo de vida es de 1 230 horas con una desviación estándar de 120 horas. ¿Hay alguna diferencia entre la media de los tiempos de vida de las dos marcas de focos a un nivel de significancia de *a*) 0.05, *b*) 0.01?

PRUEBAS EN LAS QUE INTERVIENE LA DISTRIBUCIÓN *T* DE STUDENT

7.77. La media del tiempo de vida de los focos eléctricos que producía una compañía era de 1 120 horas con una desviación estándar de 125 horas. Una muestra de 8 unidades de un suministro de focos de fabricación reciente mostró que la media del tiempo de vida es de 1 070 horas. Pruebe la hipótesis de que no ha cambiado la media del tiempo de vida de los focos a un nivel de significancia de *a*) 0.05, *b*) 0.01.

7.78. En el problema 7.77 pruebe la hipótesis $\mu = 1\ 120$ horas contra la hipótesis alternativa $\mu < 1\ 120$ horas con base en un nivel de significancia de *a*) 0.05, *b*) 0.01.

7.79. Las especificaciones para la producción de una aleación particular exigen 23.2% de cobre. Una muestra de 10 análisis del producto indica que la media del contenido de cobre es de 23.5% y una desviación estándar de 0.24%. ¿Es posible concluir que el producto cumple con las especificaciones que se requieren a un nivel de significancia de *a*) 0.01, *b*) 0.05?

7.80. En el problema 7.79, pruebe la hipótesis de que la media del contenido de cobre es mayor que el que exigen las especificaciones, con base en un nivel de significancia de *a*) 0.01, *b*) 0.05.

7.81. Un experto en eficiencia afirma que al introducir un nuevo tipo de maquinaria al proceso de producción puede disminuir de manera sustancial el tiempo que éste requiere. Por el gasto que implica el mantenimiento de las máquinas, la dirección piensa que al menos que se pueda reducir el tiempo de producción 8% como mínimo, no pueden pagar para introducir el proceso. Seis experimentos mostraron que el tiempo de producción decrece 8.4% con una desviación estándar de 0.32%. Con base en un nivel de significancia de *a*) 0.01 y *b*) 0.05, pruebe la hipótesis de que se debe introducir la nueva maquinaria.

7.82. Se probaron dos tipos de soluciones químicas, *A* y *B*, respecto de su pH (grado de acidez de la solución). Un análisis de 6 muestras de *A* mostró un pH medio de 7.52 con una desviación estándar de 0.024. Un análisis de 5 muestras de *B* mostró un pH medio de 7.49 con una desviación estándar de 0.032. Mediante el empleo de un nivel de significancia de 0.05, determine si los dos tipos de solución tienen diferentes valores de pH.

7.83. En un examen de psicología, 12 estudiantes en una clase tuvieron una calificación media de 78 con una desviación estándar de 6, mientras que los de otra clase tuvieron 74 como calificación media con una desviación estándar de 8. Usando el nivel de significancia de 0.05, determine si el primer grupo es superior al segundo.

PRUEBAS EN LAS QUE INTERVIENE LA DISTRIBUCIÓN JI-CUADRADA

7.84. La desviación estándar de las resistencias a la ruptura de ciertos cables que produce una compañía es de 240 libras. Después de introducir un cambio en el proceso de producción, las resistencias a la ruptura de una muestra de 8 cables mostraron una desviación estándar de 300 libras. Investigue la significancia del supuesto aumento de la variabilidad utilizando un nivel de significancia de *a*) 0.05, *b*) 0.01.

7.85. Se obtiene la temperatura anual de una ciudad mediante la determinación de la media de las temperaturas medias el día 15 de cada mes. La desviación estándar de las temperaturas anuales de la ciudad durante un periodo de 100 años fue de 16° Fahrenheit. Durante los últimos 15 años se calculó una desviación estándar de las temperaturas anuales de 10° Fahrenheit. Pruebe la hipótesis de que las temperaturas en la ciudad varían menos que en el pasado, empleando un nivel de significancia de *a*) 0.05, *b*) 0.01.

7.86. En el problema 7.77, una muestra de 20 focos eléctricos reveló una desviación estándar de 150 horas en su tiempo de vida. ¿Se concluiría que esto no es común? Explique su respuesta.

PRUEBAS EN LAS QUE INTERVIENE LA DISTRIBUCIÓN *F*

7.87. Dos muestras que consisten en 21 y 9 observaciones tienen varianzas dadas por $s_1^2 = 16$ y $s_2^2 = 8$, respectivamente. Pruebe la hipótesis de que la varianza de la primera población es mayor que la de la segunda a un nivel de significancia de *a*) 0.05 y *b*) 0.01.

7.88. Resuelva el problema 7.87 si las dos muestras consisten en 60 y 120 observaciones, respectivamente.

7.89. En el problema 7.82, ¿es posible concluir que hay una diferencia significativa en la variabilidad de los valores de pH de las dos soluciones a un nivel de significancia de 0.10?

CURVAS CARACTERÍSTICAS DE OPERACIÓN

7.90. Con referencia al problema 7.63, determine la probabilidad de aceptar la hipótesis de que hay proporciones iguales de canicas rojas y azules cuando la proporción real p de canicas rojas es *a*) 0.6, *b*) 0.7, *c*) 0.8, *d*) 0.9, *e*) 0.3.

7.91. Represente los resultados del problema 7.90 mediante una gráfica de *a*) β *vs.* p, *b*) $(1 - \beta)$ *vs.* p. Compare estas gráficas con las del problema 7.25 con base en la analogía respectiva entre canicas rojas y azules con caras y cruces, respectivamente.

GRÁFICAS DE CONTROL DE CALIDAD

7.92. En el pasado un tipo específico de hilo que produce un fabricante ha tenido una media de fuerza de rompimiento de 8.64 onzas y una desviación estándar de 1.28 onzas. Para determinar si el producto cumple con los estándares, cada 3 horas se toma una muestra de 16 piezas de hilo y se determina la media de la fuerza de rompimiento. Calcule los límites de control de *a*) 99.73% o 3σ, *b*) 99% y *c*) 95% sobre una gráfica de control de calidad y explique sus aplicaciones.

7.93. En promedio casi 3% de los pernos que produce una compañía son defectuosos. Para mantener este desempeño de calidad, se examina una muestra de 200 pernos producidos cada 4 horas. Determine los límites de control de *a*) 99% y *b*) 95% del número de pernos defectuosos en cada muestra. Observe que sólo se necesitan *límites de control superiores* en este caso.

AJUSTE DE DATOS MEDIANTE DISTRIBUCIONES TEÓRICAS

7.94. Ajuste una distribución binomial a los datos de la tabla 7-24.

Tabla 7-24

x	0	1	2	3	4
f	30	62	46	10	2

LA PRUEBA JI-CUADRADA

7.99. En 60 lanzamientos de una moneda se observaron 37 caras y 23 cruces. Pruebe la hipótesis de que la moneda no está cargada con base en un nivel de significancia de *a*) 0.05, *b*) 0.01.

7.100. Resuelva el problema 7.99 mediante el uso de la corrección de Yates.

7.101. Durante un largo periodo, las calificaciones dadas por un grupo de instructores de un curso mostraron los siguientes promedios: A, 12%; B, 18%; C, 40%; D, 18%, y F, 12%. Una nueva instructora proporciona en dos semestres 22 A, 34 B, 66 C, 16 D y 12 F. Determine a un nivel de significancia de 0.05 si la nueva instructora sigue el patrón de calificaciones que tienen los demás.

7.102. Se lanzaron tres monedas juntas un total de 240 veces, y cada vez se observó el número de caras que aparecía. Los resultados se muestran en la tabla 7-26 junto con los resultados esperados con base en la hipótesis de que las monedas no están cargadas. Pruebe esta hipótesis a un nivel de significancia de 0.05.

Tabla 7-26

	0 caras	1 cara	2 caras	3 caras
Frecuencia observada	24	108	95	23
Frecuencia esperada	30	90	90	30

7.103. El número de libros prestados por una biblioteca pública durante cierta semana se presenta en la tabla 7-27. Pruebe la hipótesis de que el número de libros prestados no depende del día de la semana, empleando un nivel de significancia de *a*) 0.05, *b*) 0.01.

Tabla 7-27

	Lun.	Mar.	Miér.	Juev.	Vier.
Número de libros prestados	135	108	120	114	146

7.104. Una urna contiene 6 canicas rojas y 3 blancas. Se sacan dos canicas de la urna de manera aleatoria, se anotan sus colores y se devuelven, reemplazan, a la urna. Este proceso se realiza 120 veces en total y los resultados que se obtuvieron se muestran en la tabla 7-28. *a*) Determine las frecuencias esperadas. *b*) Calcule con un nivel de significancia de 0.05 si los resultados que se obtuvieron son consistentes con los esperados.

Tabla 7-28

	0 rojas, 2 blancas	1 roja, 1 blanca	2 rojas, 0 blancas
Número de extracciones	6	53	61

7.105. Se seleccionaron 200 pernos de manera aleatoria de la producción de cada una de 4 máquinas. Los números de pernos defectuosos fueron 2, 9, 10, 3. Determine si hay una diferencia significativa entre las máquinas usando un nivel de significancia de 0.05.

BONDAD DE AJUSTE

7.106. *a*) Utilice la prueba ji-cuadrada para determinar la bondad de ajuste de los datos del problema 7.94. *b*) ¿El ajuste no es "demasiado bondadoso"? Utilice un nivel de significancia de 0.05.

7.107. Utilice la prueba ji-cuadrada para determinar la bondad de ajuste de los datos en *a*) el problema 7.95, *b*) el problema 7.96. Emplee un nivel de significancia de 0.05 y en cada caso determine si el ajuste es "no es demasiado bondadoso".

7.108. Aplique la prueba ji-cuadrada para determinar la bondad de ajuste de los datos en *a*) el problema 7.97, *b*) el problema 7.98. ¿Es el resultado de *a*) consistente con el del problema 7.106?

TABLAS DE CONTINGENCIA

7.109. La tabla 7-29 muestra el resultado de un experimento para investigar el efecto de vacunación de animales de laboratorio contra una cierta enfermedad. Utilizando un nivel de significancia de *a*) 0.01, *b*) 0.05, pruebe la hipótesis de que no hay diferencia entre los grupos vacunados y los no vacunados, esto es, que la vacunación y esta enfermedad son independientes.

Tabla 7-29

	Se enfermaron	No se enfermaron
Vacunados	9	42
No vacunados	17	28

7.110. Resuelva el problema 7.109 mediante la corrección de Yates.

7.111. La tabla 7-30 muestra los números de estudiantes en cada una de dos clases, *A* y *B*, que aprobaron y no aprobaron un examen aplicado a los dos grupos. Con un nivel de significancia de *a*) 0.05 y *b*) 0.01, pruebe la hipótesis de que no hay diferencia entre las dos clases. Resuelva el problema con y sin la corrección de Yates.

Tabla 7-30

	Aprobados	No aprobados
Clase A	72	17
Clase B	64	23

7.112. De un grupo de pacientes que se quejaron que no dormían bien, a algunos se les dieron somníferos mientras que a otros pastillas de azúcar (aunque todos *pensaban* que recibían los somníferos). Después les preguntaron si los somníferos funcionaron o no. En la tabla 7-31 se muestran los resultados de sus respuestas. Bajo el supuesto de que los pacientes dijeron la verdad, pruebe la hipótesis de que no hay diferencia entre los somníferos y las pastillas de azúcar a un nivel de significancia de 0.05.

COEFICIENTE DE CONTINGENCIA

7.116. La tabla 7-35 muestra la relación entre el color del pelo y el de los ojos en una muestra de 200 mujeres. *a*) Calcule el coeficiente de contingencia sin y con la corrección de Yates. *b*) Compare el resultado de *a*) con el coeficiente máximo de contingencia.

Tabla 7-35

		COLOR DE PELO	
		Rubio	No rubio
COLOR DE OJOS	Azules	49	25
	No azules	30	96

7.117. Calcule el coeficiente de contingencia de los datos de *a*) el problema 7.109 y *b*) el problema 7.111 sin y con la corrección de Yates.

7.118. Determine el coeficiente de contingencia de los datos del problema 7.114.

PROBLEMAS DIVERSOS

7.119. Dos urnas, *A* y *B*, contienen números iguales de canicas, pero no se conocen las proporciones de canicas rojas y blancas en cada urna. Una muestra de 50 canicas extraída con reemplazo en cada de las urnas mostró 32 canicas rojas de *A* y 23 de *B*. Con base en un nivel de significancia de 0.05, pruebe la hipótesis de que *a*) las dos urnas tienen proporciones iguales de canicas de ambos colores y *b*) que *A* tiene una proporción mayor de canicas rojas que *B*.

7.120. Con referencia al problema 7.54, determine el menor número de preguntas que un estudiante tiene que contestar correctamente para que el instructor esté seguro, a un nivel de significancia de *a*) 0.05, *b*) 0.01, *c*) 0.001, *d*) 0.06, de que el estudiante no está adivinando. Analice los resultados.

7.121. Una moneda que se lanza 8 veces resulta en caras 7 veces. ¿Es posible rechazar la hipótesis de que la moneda no está cargada a un nivel de significancia de *a*) 0.05, *b*) 0.10, *c*) 0.01? Emplee la prueba de dos colas.

7.122. El porcentaje de calificaciones A dadas en un curso de física en una universidad determinada durante un periodo largo fue de 10%. Durante un periodo particular, había 40 calificaciones A en un grupo de 300 estudiantes. Pruebe la significancia de este resultado a un nivel de *a*) 0.05, *b*) 0.01.

7.123. Utilizando la gasolina de marca A, el número medio de kilómetros por litro recorrido por 5 vehículos similares bajo condiciones idénticas fue de 22.6 con una desviación estándar de 0.48. Con la marca B, el número medio fue de 21.4 con una desviación estándar de 0.54. Elija un nivel de significancia de 0.05 e investigue si la marca A rinde mayor kilometraje que la B.

7.124. En el problema 7.123, ¿hay más variabilidad en kilómetros por litro usando la marca B, de las que hay si usa la marca A? Explique su respuesta.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

7.85. La presunta disminución es significativa a un nivel de significancia de 0.05 pero no a uno de 0.01.

7.86. Se concluiría a un nivel de 0.05 que el resultado es poco común pero no a un nivel de 0.01.

7.87. No es posible concluir a cualquiera de los dos niveles que la primera varianza es mayor que la segunda.

7.88. A ambos niveles se puede concluir que la primera varianza es mayor que la segunda.

7.89. No. **7.90.** *a*) 0.3112 *b*) 0.0118 *c*) 0 *d*) 0 *e*) 0.0118.

7.92. *a*) 8.64 ± 0.96 onzas *b*) 8.64 ± 0.83 onzas *c*) 8.64 ± 0.63 onzas **7.93.** *a*) 6 *b*) 4

7.94. $f(x) = {}_4C_x(0.32)^x(0.68)^{4-x}$; las frecuencias esperadas son 32, 60, 43, 13 y 2, respectivamente.

7.95. Las frecuencias esperadas son 1.7, 5.5, 12.0, 15.9, 13.7, 7.6, 2.7 y 0.6, respectivamente.

7.96. Las frecuencias esperadas son 1.1, 4.0, 11.1, 23.9, 39.5, 50.2, 49.0, 36.6, 21.1, 9.4, 3.1 y 1.0, respectivamente.

7.97. Las frecuencias esperadas son 41.7, 53.4, 34.2, 14.6 y 4.7, respectivamente.

7.98. $f(x) = \dfrac{(0.61)^x e^{-0.61}}{x!}$; las frecuencias esperadas son 108.7, 66.3, 20.2, 4.1 y 0.7, respectivamente.

7.99. No es posible rechazar la hipótesis a cualquier nivel.

7.100. La conclusión es igual que la anterior.

7.101. La nueva instructora no sigue el patrón de los demás. (*Quizá* el hecho de que las calificaciones casualmente sean mejores que el promedio se debe a mejor capacidad de enseñar o normas de evaluación menos exigentes, o los dos.)

7.102. No existe una razón para rechazar la hipótesis de que las monedas no están cargadas.

7.103. No hay por qué rechazar la hipótesis a cualquier nivel.

7.104. *a*) 10, 60, 50, respectivamente. *b*) No se puede rechazar la hipótesis de que los resultados son iguales a los esperados a un nivel de significancia de 0.05.

7.105. La diferencia es significativa a un nivel de 0.05.

7.106. *a*) El ajuste es bueno. *b*) No.

7.107. *a*) El ajuste es "demasiado bondadoso". *b*) El ajuste es pobre a un nivel de 0.05.

7.108. *a*) El ajuste es muy pobre a un nivel de 0.05. Puesto que la distribución binomial proporciona un buen ajuste de los datos, esto es consistente con el problema 7.109. *b*) El ajuste es bueno pero no "demasiado bondadoso".

7.109. Se puede rechazar la hipótesis a un nivel de 0.05 pero no a uno de 0.01.

7.110. La misma conclusión.

7.111. No es posible rechazar la hipótesis a cualquiera de los niveles.

7.112. No se puede rechazar la hipótesis a un nivel de 0.05.

7.113. Se puede rechazar la hipótesis a los dos niveles.

7.114. A ambos niveles es posible rechazar la hipótesis.

7.115. No es posible rechazar la hipótesis a cualquier nivel.

7.116. *a*) 0.3863, 0.3779 (con la corrección de Yates).

7.117. *a*) 0.2250, 0.1985 (corregida) *b*) 0.0872, 0.0738 corregida.

7.118. 0.4651.

7.119. *a*) Una prueba de dos colas a un nivel de 0.05 no sirve para rechazar la hipótesis de proporciones iguales.
b) Una prueba de una cola a un nivel de significancia de 0.05 indica que *A* tiene una proporción más grande de canicas rojas que *B*.

7.120. *a*) 9 *b*) 10 *c*) 10 *d*) 8

7.121. *a*) No *b*) Sí *c*) No

7.122. Utilizando una prueba de una cola, el resultado es significativo a un nivel de 0.05 pero no a un nivel de 0.01.

7.123. Es posible concluir que la marca *A* es mejor que la marca *B* a un nivel de 0.05.

7.124. No a un nivel de 0.05.

PROBLEMAS APORTADOS

1. En la elaboración de calzado industrial figura el proceso de pegar (unir) la suela con el calzado, para lo cual se utilizan adhesivos. En un día en particular el índice de rechazos es de 18% por fallas importantes en el proceso. El inspector de calidad monitorea el proceso, y decide verificar 15 zapatos para realizar la prueba de pegado. Si en dicho día se producen 200 pares (400 zapatos):

a) ¿Cuál es la probabilidad de que encuentre exactamente 6 zapatos defectuosos?
b) ¿Cuál es la probabilidad de que obtenga a lo sumo 8 zapatos defectuosos?
c) ¿Cuál es la probabilidad de detectar entre 4 y 9 zapatos defectuosos?

2. En una colonia popular se ha realizado un estudio para determinar el peso de los niños entre 3 y 5 años de edad. El estudio muestra que los resultados de los datos siguen un comportamiento aproximado a la distribución normal, que la media de las observaciones es de 17.3 kilogramos, y que la desviación estándar de dichos datos es de 4.54.
Se elige un niño al azar de este grupo estudiado. Calcule las siguientes probabilidades:

a) Que su peso esté entre 16.4 y 18.2 kilogramos.
b) Que pese más de 17.5 kilogramos.
c) Que pese menos de 17.0 kilogramos.

3. En un estacionamiento de un centro comercial existen dos accesos, y por el primero de ellos ingresan un promedio de 15.6 automóviles por hora, mientras que por el segundo el promedio es de 7.3 automóviles. ¿Cuál es la probabilidad de que en una hora en particular ingresen a dicho estacionamiento más de 22 automóviles?

CAPÍTULO 8

AJUSTE DE CURVAS

En la práctica, a menudo se encuentra que existe una relación entre dos (o más) variables, y se desea expresar esta relación en forma matemática determinando una ecuación que relacione las variables.

Un primer paso es la recopilación de datos que muestran valores que corresponden a las variables. Por ejemplo, suponga que x y y denotan, respectivamente, la estatura y peso de un hombre adulto. Entonces, una muestra de n individuos revelaría las estaturas $x_1, x_2, \ldots, x_n$ y los pesos correspondientes $y_1, y_2, \ldots, y_n$.

El siguiente paso es graficar los puntos $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ en un sistema de coordenadas rectangular. Algunas veces al conjunto de puntos resultante se le llama *diagrama de dispersión.*

A partir del diagrama de dispersión, a menudo es posible visualizar una curva lisa muy próxima a los datos. Esta curva se le denomina *curva de aproximación*. En la figura 8-1, por ejemplo, los datos parecen aproximarse bien por medio de una recta, y se observa que existe una *relación lineal* entre las variables. No obstante, en la figura 8-2, aunque existe una relación entre las variables, ésta no es lineal y, por tanto, se denomina una *relación no lineal*. En la figura 8-3 parece no haber relación entre las variables.

El problema general es encontrar ecuaciones de curvas de aproximación que ajusten a ciertos conjuntos de datos, el procedimiento se conoce como *ajuste de curvas*. En la práctica, el diagrama de dispersión sugiere el tipo de ecuación. En el caso de la figura 8-1 se podría usar una recta

$$y = a + bx \tag{1}$$

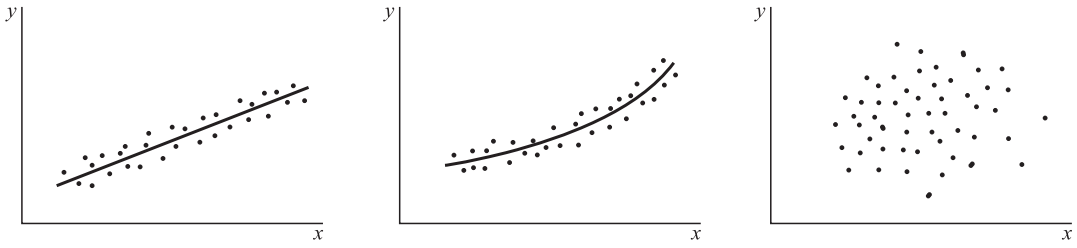
mientras que en la figura 8-2 es posible probar una *parábola* o *curva cuadrática:*

$$y = a + bx + cx^2 \tag{2}$$

Algunas veces es útil graficar los diagramas de dispersión en términos de *variables transformadas.* Por ejemplo, si log y *vs*. x conduce a una recta, se intentaría que log $y = a + bx$ como una ecuación para representar la curva de aproximación.

REGRESIÓN

Uno de los propósitos principales del ajuste de curvas es estimar una de las variables (la *variable dependiente*) a partir de otra (la *variable independiente*). El proceso de estimación se acostumbra llamarlo *regresión*. Si se va a estimar y a partir de x por medio de alguna ecuación, ésta se denomina *ecuación de regresión de y en x*, y la curva correspondiente, *curva de regresión de y en x.*



MÉTODO DE MÍNIMOS CUADRADOS

Generalmente, más de una curva de cierto tipo parecerá ajustarse al conjunto de datos. Para evitar juicios individuales cuando se construyen las rectas, parábolas u otras curvas de aproximación, es necesario concordar con la definición de la "recta de mejor ajuste" o la "parábola de mejor ajuste", etcétera.

Para motivar una posible definición, considere la figura 8-4, donde los puntos dato son $(x_1, y_1), \ldots, (x_n, y_n)$. Para un valor dado de x, digamos x_1, habrá una diferencia entre el valor y_1 y el valor correspondiente que se determina a partir de la curva C. Esta diferencia se denota d_1, que se conoce a veces como una *desviación, error* o *residuo*, y puede ser positiva, negativa o cero. De manera similar, se obtienen las desviaciones a $x_2, \ldots, x_n$, que corresponden a los valores $d_2, \ldots, d_n$.

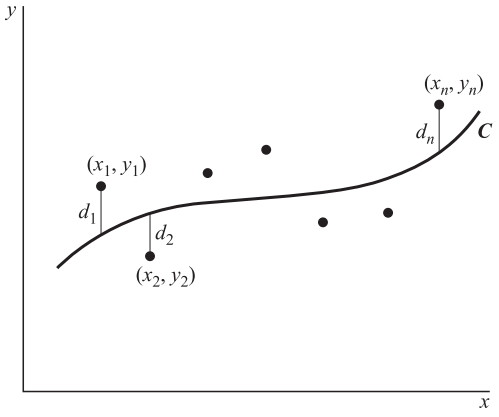


Figura 8-4

Una medida de bondad de ajuste de la curva C al conjunto de datos la proporciona la cantidad $d_1^2 + d_2^2 + \cdots + d_n^2$. Si esta cantidad es pequeña, el ajuste es bueno, si es grande, el ajuste es malo. Por tanto, a veces se tiene la siguiente definición.

Definición De todas las curvas en una familia dada de curvas de aproximación obtenida de un conjunto n de puntos de datos, una curva que tiene la propiedad de

$$d_1^2 + d_2^2 + \cdots + d_n^2 = \text{un mínimo}$$

se denomina *curva de mejor ajuste* de la familia.

Una curva que tiene esta propiedad es ajustar los datos utilizando *mínimos cuadrados* se le llama *curva de regresión de mínimos cuadrados* o, simplemente, *curva de mínimos cuadrados*. A una recta que tiene estas propiedades se le conoce como *recta de mínimos cuadrados*; una parábola con esta propiedad se llama *parábola de mínimos cuadrados*, etcétera.

Se acostumbra emplear la definición anterior cuando x es la variable independiente y y es la variable dependiente. Si x es la variable dependiente, se modifica la definición, pues se consideran las desviaciones horizontales en lugar de las verticales, lo cual explica el intercambio de los ejes x y y. Estas dos definiciones llevan, en general, a dos diferentes curvas de mínimos cuadrados. A menos que se especifique lo contrario, se considerará a y como la variable dependiente y a x como la variable independiente.

Es posible definir otra curva de mínimos cuadrados considerando las distancias perpendiculares desde los puntos datos a la curva en vez de distancias horizontales o verticales. Sin embargo, este enfoque no se usa con frecuencia.

RECTA DE MÍNIMOS CUADRADOS

Al usar la definición anterior, se puede demostrar (vea el problema 8.3) que la recta de mínimos cuadrados que mejor se aproxima al conjunto de puntos $(x_1, y_1), \ldots, (x_n, y_n)$ tiene la ecuación

$$y = a + bx \tag{3}$$

donde las constantes a y b se determinan al resolver simultáneamente las ecuaciones

$$\begin{aligned} \sum y &= an + b\sum x \\ \sum xy &= a\sum x + b\sum x^2 \end{aligned} \tag{4}$$

que se denominan ecuaciones normales para la recta de mínimos cuadrados. Note que se ha usado por brevedad $\sum y$, $\sum xy$ en vez de $\sum_{j=1}^{n} y_j$, $\sum_{j=1}^{n} x_j y_j$. Las ecuaciones normales (4) se recuerdan con facilidad si se observa que la primera ecuación se obtiene formalmente mediante la suma de ambos lados de (3), mientras que la segunda ecuación se obtiene formalmente multiplicando primero ambos lados de (3) por x y después realizar la suma. Es evidente que esto no es una deducción de las ecuaciones normales sino sólo un medio para recordarlas.

Los valores de a y b que se obtuvieron en (4) están dados por

$$a = \frac{\left(\sum y\right)\left(\sum x^2\right) - \left(\sum x\right)\left(\sum xy\right)}{n\sum x^2 - \left(\sum x\right)^2} \qquad b = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} \tag{5}$$

El resultado de b en (5) se expresa como

$$b = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2} \tag{6}$$

Aquí, como es usual, una barra indica *media*, esto es, $\bar{x} = (\sum x)/n$. Al dividir ambos lados de la primera ecuación normal (4) entre n, se produce

$$\bar{y} = a + b\bar{x} \tag{7}$$

Si se desea, es posible calcular primero b a partir de (5) o (6) y después usar (7) para determinar $a = \bar{y} - b\bar{x}$. Esto equivale a escribir la recta de mínimos cuadrados como

$$y - \bar{y} = b(x - \bar{x}) \qquad \text{o} \qquad y - \bar{y} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}(x - \bar{x}) \tag{8}$$

El resultado (8) muestra que la constante b, que es la *pendiente* de la recta (3), es la constante fundamental para determinar la recta. A partir de (8) también se observa que la recta de mínimos cuadrados pasa por el punto $(\bar{x}, \bar{y})$, que se llama *centroide* o *centro de gravedad* de los datos.

La pendiente b de la recta de regresión es independiente del origen de coordenadas. Esto significa que si se hace la transformación (que a menudo se llama *traslación de ejes*) dada por

$$x = x' + h \qquad y = y' + k \tag{9}$$

donde h y k son constantes cualesquiera, entonces b también está dada por

$$b = \frac{n\sum x'y' - \left(\sum x'\right)\left(\sum y'\right)}{n\sum x'^2 - \left(\sum x'\right)^2} = \frac{\sum (x' - \overline{x'})(y' - \overline{y'})}{\sum (x' - \overline{x'})^2} \tag{10}$$

$$= \frac{\sum x'y'}{\sum x'^2} \tag{11}$$

Los resultados (10) y (11) a menudo son útiles para simplificar la labor que implica obtener la recta de mínimos cuadrados.

Los comentarios anteriores también se cumplen en el caso de la recta de regresión x en y. Los resultados se obtienen formalmente mediante el intercambio de x y y. Por ejemplo, la recta de regresión de mínimos cuadrados de x en y es

$$x - \bar{x} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (y - \bar{y})^2}(y - \bar{y}) \tag{12}$$

En general, se considera que (12) no es la misma recta que (8).

Recta de mínimos cuadrados en términos de varianzas y covarianza muestrales

Las varianzas y la covarianza muestrales de x y y están dadas por

$$s_x^2 = \frac{\sum (x - \bar{x})^2}{n}, \qquad s_y^2 = \frac{\sum (y - \bar{y})^2}{n}, \qquad s_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n} \tag{13}$$

En términos de éstas, las rectas de regresión de mínimos cuadrados de y en x y de x en y se pueden escribir, respectivamente, como

$$y - \bar{y} = \frac{s_{xy}}{s_x^2}(x - \bar{x}) \qquad \text{y} \qquad x - \bar{x} = \frac{s_{xy}}{s_y^2}(y - \bar{y}) \tag{14}$$

si formalmente se define el *coeficiente de correlación muestral* mediante [comparar (54), página 82]

$$r = \frac{s_{xy}}{s_x s_y} \tag{15}$$

entonces (14) se escribe como

$$\frac{y - \bar{y}}{s_y} = r\left(\frac{x - \bar{x}}{s_x}\right) \qquad \text{y} \qquad \frac{x - \bar{x}}{s_x} = r\left(\frac{y - \bar{y}}{s_y}\right) \tag{16}$$

Dado el hecho de que $(x - \bar{x})/s_x$ y $(y - \bar{y})/s_y$ son valores muestrales estandarizados o puntajes estándar, los resultados en (16) proporcionan una manera muy sencilla de recordar las rectas de regresión. Es evidente que las dos rectas en (16) son diferentes, a menos que $r = \pm 1$, en cuyo caso todos los puntos muestrales se ubican en la recta [esto se observa en (26)] y existe *correlación y regresión lineal perfecta.*

Es también de interés observar que si se escriben las dos rectas de regresión (16) como $y = a + bx$, $x = c + dy$, respectivamente, entonces

$$bd = r^2 \tag{17}$$

Hasta ahora no se ha considerado el significado preciso del coeficiente de correlación sino que sólo se ha definido formalmente en términos de las varianzas y covarianzas. En la página 270 se dará el significado.

Parábola de mínimos cuadrados

Las ideas anteriores se extienden fácilmente. Por ejemplo, la *parábola de mínimos cuadrados* que ajusta al conjunto de puntos muestrales está dada por

$$y = a + bx + cx^2 \tag{18}$$

donde a, b y c se determinan mediante las *ecuaciones normales*

$$\begin{aligned} \sum y &= na + b\sum x + c\sum x^2 \\ \sum xy &= a\sum x + b\sum x^2 + c\sum x^3 \\ \sum x^2 y &= a\sum x^2 + b\sum x^3 + c\sum x^4 \end{aligned} \tag{19}$$

Estas ecuaciones se obtienen de manera formal mediante la suma de ambos lados de (18), y después se multiplican de manera sucesiva por 1, x y x^2, respectivamente.

REGRESIÓN MÚLTIPLE

También es posible generalizar las ideas anteriores a más variables. Por ejemplo, si se siente que hay una relación lineal entre una variable dependiente z y dos variables independientes, x y y, entonces se busca una ecuación que relaciona las variables que tienen la forma

$$z = a + bx + cy \tag{20}$$

Ésta se llama *ecuación de regresión de z en x* y *y*. Si x es la variable dependiente, una ecuación similar se llamaría *ecuación de regresión de x en y y z*.

Puesto que (20) representa un plano en un sistema de coordenadas rectangular tridimensional, a menudo se denomina *plano de regresión*. Para encontrar el plano de regresión de mínimos cuadrados, se determinan a, b, c en (20), de modo que

$$\begin{aligned} \sum z &= na + b\sum x + c\sum y \\ \sum xz &= a\sum x + b\sum x^2 + c\sum xy \\ \sum yz &= a\sum y + b\sum xy + c\sum y^2 \end{aligned} \tag{21}$$

Estas ecuaciones, llamadas *ecuaciones normales* correspondientes a (20), se obtienen como resultado de aplicar una definición similar a la de la página 266. Observe que se obtienen formalmente de (20) multiplicando por 1, x, y, respectivamente, y sumando.

Es fácil generalizar a más variables que implican ecuaciones lineales o no lineales que llevan a *superficies de regresión* en espacios tridimensionales o superiores.

ERROR ESTÁNDAR DE ESTIMACIONES

Si se deja que y_{est} denote el valor estimado de y para un valor dado de x, como se obtiene en la curva de regresión y en x, entonces la cantidad

$$s_{y.x} = \sqrt{\frac{\sum (y - y_{est})^2}{n}} \tag{22}$$

que se llama *error estándar de estimación de y en x,* proporciona una medida de la dispersión alrededor de la curva de regresión. Puesto que $\sum (y - y_{est})^2 = \sum d^2$, como se usa en la definición de la página 266, se observa que de todas las posibles curvas de regresión, la curva de mínimos cuadrados tiene el error estándar de estimación más pequeño.

En el caso de una recta de regresión $y_{est} = a + bx$, donde a y b están dadas por (4), se tiene

$$s_{y.x}^2 = \frac{\sum y^2 - a\sum y - b\sum xy}{n} \tag{23}$$

o

$$s_{y.x}^2 = \frac{\sum (y - \bar{y})^2 - b\sum (x - \bar{x})(y - \bar{y})}{n} \tag{24}$$

También se expresa $s_{y.x}^2$ para la recta de mínimos cuadrados en términos del coeficiente de varianza y correlación como

$$s_{y.x}^2 = s_y^2(1 - r^2) \tag{25}$$

a partir del cual se concluye como un corolario que $r^2 \leq 1$, esto es, $-1 \leq r \leq 1$.

$= ns^2/(n-1)$,
hay una estimación insesgada del cuadrado del error estándar de estimación que está dada por $\hat{s}^2_{y.x} = ns^2_{y.x}/(n-2)$. Por este motivo, algunos estadísticos prefieren indicar (22) con $n-2$ en vez de n en el denominador.

Es fácil adecuar los comentarios anteriores para la recta de regresión de x en y (en cuyo caso el error estándar de la estimación se denota por medio de $s_{y.x}$) o para regresión múltiple o no lineal.

COEFICIENTE DE CORRELACIÓN LINEAL

Hasta el momento se ha definido el coeficiente de correlación formalmente por (15); sin embargo, no se ha examinado su significancia. En el intento de hacer esto, debemos señalar que a partir de (25) y de las definiciones de $s_{y.x}$ y s_y, se tiene

$$r^2 = 1 - \frac{\sum(y - y_{\text{est}})^2}{\sum(y - \bar{y})^2} \tag{26}$$

Ahora es posible demostrar que (vea el problema 8.24)

$$\sum(y - \bar{y})^2 = \sum(y - y_{\text{est}})^2 + \sum(y_{\text{est}} - \bar{y})^2 \tag{27}$$

La cantidad a la izquierda de (27) es llamada *variación total*. Luego, la primera suma a la derecha de (27) se denomina *variación inexplicable* y la segunda suma, *variación explicada*. Esta terminología surge porque las desviaciones $y - y_{\text{est}}$ se comportan de manera aleatoria o no predictiva, mientras que las desviaciones $y_{\text{est}} - \bar{y}$ se explican mediante la recta de regresión de mínimos cuadrados y, por ello, tienden a seguir un patrón definido. Se deduce de (26) y (27) que

$$r^2 = \frac{\sum(y_{\text{est}} - \bar{y})^2}{\sum(y - \bar{y})^2} = \frac{\text{variación explicada}}{\text{variación total}} \tag{28}$$

Por tanto, se interpreta r^2 como la porción de la variación total que se explica por medio de la recta de regresión de mínimos cuadrados. En otras palabras, r mide *qué tan bien* la recta de regresión de mínimos cuadrados ajusta los datos muestrales. Si la recta de regresión explica *completamente* la variación total, esto es, si $r^2 = 1$, es decir, $r = \pm 1$, existe una *correlación lineal perfecta* (y en tal caso también *regresión lineal perfecta*). Por otro lado, si no se explica en forma completa la variación total, entonces la variación explicada es cero y por ello $r = 0$. En la práctica, la cantidad r^2, que algunas veces se le llama *coeficiente de determinación*, se ubica entre 0 y 1.

El coeficiente de correlación se calcula a partir de los resultados de

$$r = \frac{s_{xy}}{s_x s_y} = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2}\sqrt{\sum(y - \bar{y})^2}} \tag{29}$$

o de

$$r^2 = \frac{\text{variación explicada}}{\text{variación total}} = \frac{\sum(y_{\text{est}} - \bar{y})^2}{\sum(y - \bar{y})^2} \tag{30}$$

que para el caso de regresión lineal son equivalentes. La fórmula (29) a menudo se denomina *fórmula de momento-producto* para correlación lineal.

Las siguientes son fórmulas equivalentes a las anteriores que se emplean a menudo en la práctica,

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}} \tag{31}$$

y

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{(\bar{x}^2 - \bar{x}^2)(\bar{y}^2 - \bar{y}^2)}} \tag{32}$$

Si se utiliza la transformación (9), página 267, se encuentra que

$$r = \frac{n\sum x'y' - \left(\sum x'\right)\left(\sum y'\right)}{\sqrt{\left[n\sum x'^2 - \left(\sum x'\right)^2\right]\left[n\sum y'^2 - \left(\sum y'\right)^2\right]}} \tag{33}$$

que indica que r es invariante bajo la traslación de ejes. En particular, si $h = \bar{x}$, $k = \bar{y}$, (33) se convierte en

$$r = \frac{\sum x'y'}{\sqrt{\left(\sum x'^2\right)\left(\sum y'^2\right)}} \tag{34}$$

que se usa a menudo en los cálculos.

El coeficiente de correlación lineal puede ser positivo o negativo. Si r es positivo, y tiende a *aumentar* con x (la pendiente de la recta de mínimos cuadrados es positiva), mientras que si r es negativo, y tiende a *decrecer* con x (la pendiente es negativa). El signo se toma en cuenta *automáticamente* si se usa el resultado (29), (31), (32), (33) o (34). Sin embargo, si se emplea (30) para obtener r, es necesario aplicar el signo adecuado.

COEFICIENTE DE CORRELACIÓN GENERALIZADO

La definición (29) [o cualquiera de las formas equivalentes de (31) a (34)] para el coeficiente de correlación implica sólo valores muestrales x, y. En consecuencia, produce el mismo número para todas las formas de curvas de regresión y es inútil como una medida de ajuste, excepto en el caso de regresión lineal, donde resulta que coincide con (30). Sin embargo, esta última definición,

$$r^2 = \frac{\text{variación explicada}}{\text{variación total}} = \frac{\sum(y_{\text{est}} - \bar{y})^2}{\sum(y - \bar{y})^2} \tag{35}$$

refleja la forma de la curva de regresión (por medio de la y_{est}) y, por tanto, es adecuada como definición del *coeficiente de correlación generalizado r*. Se recurre a (35) para obtener coeficientes de correlación no lineales (que miden qué tan bien ajusta los datos una *curva de regresión no lineal*) o, mediante una generalización adecuada, *coeficientes de correlación múltiple*. La conexión (25) entre el coeficiente de correlación y el error estándar de estimación se mantiene también para correlación no lineal.

Puesto que un coeficiente de correlación sólo mide qué tan bien ajusta cierta curva de regresión (o superficie) los datos muestrales, es evidente que no tiene sentido emplear el coeficiente de correlación lineal donde los datos no son lineales. Se supone, sin embargo, que se aplica (29) a datos no lineales y se obtiene un valor que es numéricamente menor que 1. Entonces, no es posible concluir que existe *poca correlación* (una conclusión a la que a veces arriban personas no familiarizadas con los fundamentos de la teoría de correlación), sino que existe *poca correlación lineal*. En realidad, puede existir una *gran* correlación no lineal.

CORRELACIÓN DE RANGOS

En lugar de emplear valores muestrales precisos, o cuando no se logra precisión, los datos se ordenan de acuerdo con el tamaño, importancia, etc., usando los números 1, 2, . . . , n. Si se clasifican de tal manera dos conjuntos de valores correspondientes, x y y, el *coeficiente de correlación de rango*, que se denota por r_{rango}, o de manera breve, r (vea el problema 8.36) está dado por

$$r_{\text{rango}} = 1 - \frac{6\sum d^2}{n(n^2 - 1)} \tag{36}$$

donde d = diferencias entre rangos correspondientes a x y y

n = número de pares de valores (x, y) en los datos.

La cantidad r_{rango} en (36) se conoce como *coeficiente de correlación de rangos de Spearman*.

INTERPRETACIÓN PROBABILÍSTICA DE LA REGRESIÓN

Un diagrama de dispersión, como el de la figura 8-1, es una representación gráfica de puntos (datos) para una muestra particular. Mediante la elección de una muestra diferente, o ampliando la original, en general se obtiene un diagrama de dispersión algo diferente. En consecuencia, cada diagrama de dispersión resultaría en una recta o curva de regresión diferente, aunque se esperaría que no difirieran una de otra de manera significativa si se toman las muestras de la misma población.

El concepto de ajuste de curvas en muestras lleva al ajuste de curva de la población de la cual se toman muestras. La dispersión de puntos alrededor de la recta o curva de regresión indica que, para cierto valor de x, existen varios valores de y *distribuidos* alrededor de la recta o curva. Esta idea de distribución lleva naturalmente al reconocimiento de que existe una conexión entre ajuste de curvas y probabilidad.

La conexión se establece al introducir las variables aleatorias X y Y, que pueden tomar los diversos valores muestrales x y y, respectivamente. Por ejemplo, X y Y pueden representar estaturas y pesos de hombres adultos en una población de la cual se toman muestras. Entonces, se supone que X y Y tienen una función de probabilidad o función de densidad conjunta $f(x, y)$, de acuerdo con el hecho de que se consideren discretas o continuas.

Considerando la función de densidad o de probabilidad conjunta $f(x, y)$, de dos variables aleatorias X y Y, de acuerdo con lo anterior es natural preguntar si existe una función $g(X)$ tal que

$$E\{[Y - g(X)]^2\} = \text{un mínimo} \tag{37}$$

Una curva con ecuación $y = g(x)$ con la propiedad (37) se llama *curva de regresión de mínimos cuadrados de Y en X*. Se tiene el siguiente teorema:

Teorema 8-1 Si X y Y son variables aleatorias con una función de densidad conjunta $f(x, y)$, entonces existe una curva de regresión de mínimos cuadrados de Y en X, con la propiedad (37), dada por

$$y = g(x) = E(Y \mid X = x) \tag{38}$$

siempre que X y Y tengan cada una varianza finita.

Observe que $E(Y \mid X = x)$ es la esperanza condicional de Y dado $X = x$, según se definió en la página 82.

Comentarios similares son posibles en cuanto a la *curva de regresión de mínimos cuadrados de X en Y*. En ese caso, (37) se sustituye por

$$E\{[X - h(Y)]^2\} = \text{un mínimo}$$

y (38) se sustituye por $x = h(y) = E(X \mid Y = y)$. Las dos curvas de regresión $y = g(x)$ y $x = h(y)$ son, en general, diferentes.

Un caso interesante surge cuando la distribución conjunta es la distribución normal bivariada dada por (49), página 117. Entonces, se tiene el siguiente teorema:

Teorema 8-2 Si X y Y son variables aleatorias con una distribución normal bivariada, entonces la curva de regresión de mínimos cuadrados de Y en X es una *recta de regresión* dada por

$$\frac{y - \mu_Y}{\sigma_Y} = \rho\left(\frac{x - \mu_X}{\sigma_X}\right) \tag{39}$$

donde

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{40}$$

representa el coeficiente de correlación poblacional.

También es posible expresar (39) como

$$y - \mu_Y = \beta(x - \mu_X) \tag{41}$$

donde

$$\beta = \frac{\sigma_{XY}}{\sigma_X^2} \tag{42}$$

Pueden hacerse comentarios similares sobre la curva de regresión de mínimos cuadrados de X en Y, la cual también resulta ser una recta [dada por (39) con X y Y, x y y, intercambiadas]. Se deben comparar estos resultados con los de la página 268.

En el caso que $f(x, y)$ no se conoce, todavía se puede usar el criterio (37) para obtener curvas de regresión aproximadas para la población. Por ejemplo, si se supone que $g(x) = \alpha + \beta x$, se obtiene la recta de regresión de mínimos cuadrados (39), donde α y β están dadas en términos de los parámetros (no conocidos) μ_X, μ_Y, σ_X, σ_Y, ρ. De manera similar, si $g(x) = \alpha + \beta x + \gamma x^2$, se obtiene la parábola de regresión de mínimos cuadrados, etc. Vea el problema 8.39.

En general, es fácil extender a la población todos los comentarios de las páginas 266 a 271 para muestras. Por ejemplo, el error estándar de estimación en el caso de una población está dado en términos del coeficiente de varianza y correlación por

$$\sigma_{Y.X}^2 = \sigma_Y^2(1 - \rho^2) \tag{43}$$

la que se debe comparar con (25), página 269.

INTERPRETACIÓN PROBABILÍSTICA DE CORRELACIÓN

De los comentarios anteriores es evidente que un coeficiente de correlación poblacional debe proporcionar una medida de la precisión con que una curva de regresión poblacional ajusta los datos de población. Todos los comentarios anteriores sobre la correlación en una muestra se aplican también a la población. Por ejemplo, si se determina $g(x)$ mediante (37), entonces

$$E[(Y - \bar{Y})^2] = E[(Y - Y_{\text{est}})^2] + E[(Y_{\text{est}} - \bar{Y})^2] \tag{44}$$

donde $Y_{\text{est}} = g(X)$ y $\bar{Y} = E(Y)$. Las tres cantidades en (44) se llaman *variaciones totales inexplicables*, y *explicadas*, respectivamente. Lo anterior conduce a la definición del *coeficiente de correlación poblacional* ρ, donde

$$\rho^2 = \frac{\text{variación explicada}}{\text{variación total}} = \frac{E[(Y_{\text{est}} - \bar{Y})^2]}{E[(Y - \bar{Y})^2]} \tag{45}$$

En el caso lineal, esto se reduce a (40). También se escriben resultados similares de (31) hasta (34) en el caso de una población y regresión lineal. Además, el resultado (45) se usa para definir ρ, en el caso no lineal.

TEORÍA MUESTRAL DE LA REGRESIÓN

La ecuación de regresión $y = a + bx$ se obtiene a partir de datos muestrales. A menudo es de interés la ecuación de regresión $y = \alpha + \beta x$ que corresponde a la población de la cual se tomó la muestra. A continuación se presentan algunas pruebas que implican una población normal. Para simplificar la notación, se acostumbra indicar los valores de variables aleatorias muestrales en vez de las variables aleatorias en sí.

1. **PRUEBA DE LA HIPÓTESIS $\beta = b$.** Para probar la hipótesis de que el coeficiente de regresión β es igual a algún valor específico b, se usa el hecho de que el estadístico

$$t = \frac{\beta - b}{s_{y.x}/s_x}\sqrt{n - 2} \tag{46}$$

tiene una distribución de Student con $n - 2$ grados de libertad. También es posible usar lo anterior para determinar intervalos de confianza de los coeficientes de regresión poblacional de los valores de la muestra. Vea los problemas 8.43 y 8.44.

2. **PRUEBA DE LA HIPÓTESIS DE VALORES PREDICTIVOS.** Sea y_0 el valor predictivo de y correspondiente a $x = x_0$ según se estima a partir de la ecuación de regresión muestral, esto es, $y_0 = a + bx_0$. Sea y_p el valor predictivo de y correspondiente a $x = x_0$ de la población. Entonces, el estadístico

$$t = \frac{(y_0 - y_p)\sqrt{n - 2}}{s_{y.x}\sqrt{n + 1 + [n(x_0 - \bar{x})^2/s_x^2]}} \tag{47}$$

tiene la distribución de Student con $n - 2$ grados de libertad. A partir de lo anterior se encuentran los límites de confianza para los valores poblacionales predictivos. Vea el problema 8.45.

$$= \frac{(y_0 - \bar{y}_p)\sqrt{n-2}}{s_{y.x}\sqrt{1 + [(x_0 - \bar{x})^2/s_x^2]}} \tag{48}$$

tiene la distribución de Student con $n - 2$ grados de libertad. Con esto se pueden encontrar los límites de confianza para los valores medios predictivos de la población. Vea el problema 8.46.

TEORÍA MUESTRAL DE CORRELACIÓN

Es común tener que estimar el coeficiente de correlación poblacional ρ a partir del coeficiente de correlación muestral r o probar las hipótesis que implican ρ. Para este propósito es necesario conocer la distribución muestral de r. En el caso $\rho = 0$, esta distribución es simétrica y es posible usar un estadístico que tiene la distribución de Student. Para $\rho \neq 0$, la distribución está sesgada. En este caso una transformación de Fisher produce un estadístico con distribución aproximadamente normal. Las siguientes pruebas resumen los procedimientos relacionados.

1. **PRUEBA DE LA HIPÓTESIS $\rho = 0$.** Aquí se usa el hecho de que el estadístico

$$t = \frac{r\sqrt{n-2}}{\sqrt{1 - r^2}} \tag{49}$$

tiene la distribución de Student con $n - 2$ grados de libertad. Vea los problemas 8.47 y 8.48.

2. **PRUEBA DE LA HIPÓTESIS $\rho = \rho_0 \neq 0$.** En este caso se recurre al hecho de que el estadístico

$$Z = \frac{1}{2}\ln\left(\frac{1 + r}{1 - r}\right) = 1.1513 \log_{10}\left(\frac{1 + r}{1 - r}\right) \tag{50}$$

se distribuye casi de manera aproximadamente normal con media y desviación estándar dadas por

$$\mu_z = \frac{1}{2}\ln\left(\frac{1 + \rho_0}{1 - \rho_0}\right) = 1.1513 \log_{10}\left(\frac{1 + \rho_0}{1 - \rho_0}\right), \qquad \sigma_Z = \frac{1}{\sqrt{n-3}} \tag{51}$$

Estos hechos también se usan para calcular los límites de confianza de los coeficientes de correlación. Vea los problemas 8.49 y 8.50. La transformación (50) se llama *transformación Z de Fisher*.

3. **SIGNIFICANCIA DE UNA DIFERENCIA ENTRE COEFICIENTES DE CORRELACIÓN.** Para determinar si dos coeficientes de correlación r_1 y r_2 que se toman de muestras de tamaños n_1 y n_2, respectivamente, difieren significativamente uno del otro, se calculan Z_1 y Z_2 correspondientes a r_1 y r_2 utilizando (50). Después se usa el hecho de que el estadístico de prueba

$$z = \frac{Z_1 - Z_2 - \mu_{Z_1 - Z_2}}{\sigma_{Z_1 - Z_2}} \tag{52}$$

donde $$\mu_{Z_1 - Z_2} = \mu_{Z_1} - \mu_{Z_2}, \qquad \sigma_{Z_1 - Z_2} = \sqrt{\sigma_{Z_1}^2 + \sigma_{Z_2}^2} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} \tag{53}$$

está distribuido normalmente. Vea el problema 8.51.

CORRELACIÓN Y DEPENDENCIA

Siempre que dos variables aleatorias X y Y tienen un coeficiente de correlación ρ diferente de cero, se sabe (teorema 3-15, página 81) que son *dependientes* en el sentido de la probabilidad (esto es, su distribución conjunta no es un factor en sus distribuciones marginales). Además, cuando $\rho \neq 0$, es posible usar una ecuación de la forma (39) para *predecir* el valor de Y a partir del valor de X.

Es importante comprender que los términos "correlación" y "dependencia" en el sentido anterior no necesariamente implican una interdependencia causal directa de X y Y, lo cual se muestra en los siguientes ejemplos.

EJEMPLO 8.1 Sean X y Y variables aleatorias que representan estaturas y pesos de individuos. Aquí existe una interdependencia directa entre X y Y.

EJEMPLO 8.2 Si X representa salarios de maestros a lo largo de los años, mientras que Y representa la cantidad de crímenes, el coeficiente de correlación puede ser diferente de cero y quizá sea posible encontrar una ecuación de regresión que permita predecir una variable a partir de la otra. Sin embargo, sería difícil decir que existe una interdependencia directa entre X y Y.

PROBLEMAS RESUELTOS

RECTA DE MÍNIMOS CUADRADOS

8.1. Una recta pasa por los puntos (x_1, y_1) y (x_2, y_2). Demostrar que la ecuación de la recta es

$$y - y_1 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1)$$

La ecuación de una recta es $y = a + bx$. Entonces, puesto que (x_1, y_1) y (x_2, y_2) son puntos de la recta, tenemos

$$y_1 = a + bx_1, \qquad y_2 = a + bx_2$$

Por tanto,

(1) $$y - y_1 = (a + bx) - (a + bx_1) = b(x - x_1)$$

(2) $$y_2 - y_1 = (a + bx_2) - (a + bx_1) = b(x_2 - x_1)$$

Si se obtiene $b = (y_2 - y_1)/(x_2 - x_2)$ a partir de (2) y sustituyendo en (1), se consigue el resultado que buscamos.

En la figura 8-5 se muestra la gráfica de la recta PQ. La constante $b = (y_2 - y_1)/(x_2 - x_2)$ es la pendiente de la recta.

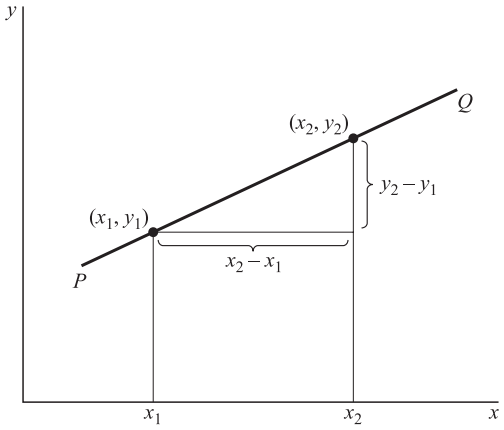


Figura 8-5

8.2. *a*) Construir una recta que se aproxime a los datos de la tabla 8-1. *b*) Encontrar la ecuación de esta recta.

Tabla 8-1

x	1	3	4	6	8	9	11	14
y	1	2	4	4	5	7	8	9

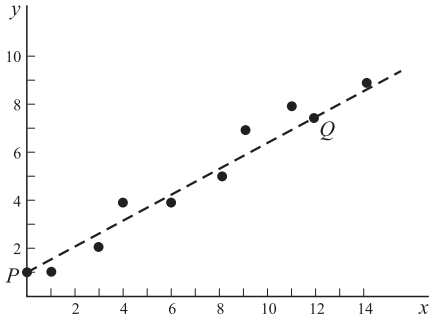


Figura 8-6

$$- \quad = \frac{7.5 - 1}{12 - 0}(x - 0)$$

o sea $y - 1 = 0.542x$ o $y = 1 + 0.542x$.

8.3. Deducir las ecuaciones normales (4), página 267, para la recta de mínimos cuadrados.

Consideremos la figura 8-7. Los valores de y en la recta de mínimos cuadrados correspondientes a $x_1, x_2, \ldots, x_n$ son

$$a + bx_1, \qquad a + bx_2, \qquad \ldots, \qquad a + bx_n$$

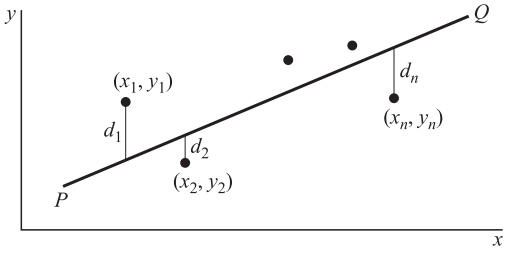


Figura 8-7

Las desviaciones verticales correspondientes son

$$d_1 = a + bx_1 - y_1, \qquad d_2 = a + bx_2 - y_2, \qquad \ldots, \qquad d_n = a + bx_n - y_n$$

Entonces, la suma de los cuadrados de las desviaciones es

$$d_1^2 + d_2^2 + \cdots + d_n^2 = (a + bx_1 - y_1)^2 + (a + bx_2 - y_2)^2 + \cdots + (a + bx_n - y_n)^2$$

o bien

$$\sum d^2 = \sum (a + bx - y)^2$$

Ésta es una función de a y b, es decir, $F(a, b) = \sum (a + bx - y)^2$. Una condición necesaria para que esto sea un mínimo (o un máximo) es que $\partial F/\partial a = 0$, $\partial F/\partial b = 0$. Puesto que

$$\frac{\partial F}{\partial a} = \sum \frac{\partial}{\partial a}(a + bx - y)^2 = \sum 2(a + bx - y)$$

$$\frac{\partial F}{\partial b} = \sum \frac{\partial}{\partial b}(a + bx - y)^2 = \sum 2x(a + bx - y)$$

obtenemos

$$\sum (a + bx - y) = 0 \qquad \sum x(a + bx - y) = 0$$

esto es,

$$\sum y = an + b\sum x \qquad \sum xy = a\sum x + b\sum x^2$$

como se requiere. Se puede demostrar que lo anterior produce un mínimo.

8.4. Ajustar la recta de mínimos cuadrados a los datos del problema 8.2 utilizando *a*) x como la variable independiente, *b*) x como la variable dependiente.

a) La ecuación de la recta es $y = a + bx$. Las ecuaciones normales son

$$\sum y = an + b\sum x$$

$$\sum xy = a\sum x + b\sum x^2$$

$$8a + 56b = 40$$
$$+ 524b = 364$$

Resolviendo simultáneamente, $a = \frac{6}{11}$ o sea 0.545, $b = \frac{7}{11}$ es decir, 0.636; y la recta de mínimos cuadrados que se requiere es $y = \frac{6}{11} + \frac{7}{11}x$ o sea $y = 0.545 + 0.636x$. Observemos que esta recta no es la que se obtuvo en el problema 8.2 usando el método manual.

Tabla 8-2

x	y	x^2	xy	y^2
1	1	1	1	1
3	2	9	6	4
4	4	16	16	16
6	4	36	24	16
8	5	64	40	25
9	7	81	63	49
11	8	121	88	64
14	9	196	126	81
$\sum x = 56$	$\sum y = 40$	$\sum x^2 = 524$	$\sum xy = 364$	$\sum y^2 = 256$

Otro método

$$a = \frac{\left(\sum y\right)\left(\sum x^2\right) - \left(\sum x\right)\left(\sum xy\right)}{n\sum x^2 - \left(\sum x\right)^2} = \frac{(40)(524) - (56)(364)}{(8)(524) - (56)^2} = \frac{6}{11} \quad \text{o sea } 0.545$$

$$b = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} = \frac{(8)(364) - (56)(40)}{(8)(524) - (56)^2} = \frac{7}{11} \quad \text{o sea } 0.636$$

b) Si se considera x como la variable dependiente y a y como la independiente, la ecuación de la recta de mínimos cuadrados es $x = c + dy$ y las ecuaciones normales son

$$\sum x = cn + d\sum y$$
$$\sum xy = c\sum y + d\sum y^2$$

Entonces, usando la tabla 8-2, las ecuaciones normales se convierten en

$$8c + 40d = 56$$
$$40c + 256d = 364$$

a partir de las cuales $c = -\frac{1}{2}$ o sea -0.50, $d = \frac{3}{2}$ o sea 1.50.

También se obtienen estos valores de

$$c = \frac{\left(\sum x\right)\left(\sum y^2\right) - \left(\sum y\right)\left(\sum xy\right)}{n\sum y^2 - \left(\sum y\right)^2} = \frac{(56)(256) - (40)(364)}{(8)(256) - (40)^2} = -0.50$$

$$d = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum y^2 - \left(\sum y\right)^2} = \frac{(8)(364) - (56)(40)}{(8)(256) - (40)^2} = 1.50$$

$y = 0.545 + 0.636x$ $= -0.500 + 1.50y$. Observemos que ambas rectas, en este caso, son prácticamente coincidentes, lo que indica que se describen muy bien los datos mediante una relación lineal.

La recta que se obtuvo en el inciso *a*) a menudo se denomina *recta de regresión de y en x* y se utiliza para estimar *y* para valores dados de *x*. La recta que se obtuvo en el inciso *b*) se llama *recta de regresión de x en y* y se emplea para estimar *x* en función de valores dados de *y*.

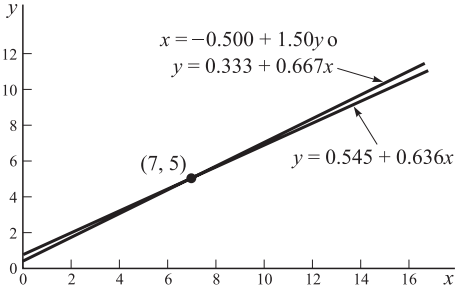


Figura 8-8

8.6. *a*) Demostrar que las dos rectas de mínimos cuadrados del problema 8.4 se intersecan en el punto $(\bar{x}, \bar{y})$. *b*) Estimar el valor de *y* cuando $x = 12$. *c*) Estimar el valor de *x* cuando $y = 3$.

$$\bar{x} = \frac{\sum x}{n} = \frac{56}{8} = 7, \qquad \bar{y} = \frac{\sum y}{n} = \frac{40}{8} = 5$$

Entonces, el punto $(\bar{x}, \bar{y})$, llamado *centroide*, es (7, 5).

a) El punto (7, 5) se ubica en la recta $y = 0.545 + 0.636x$ o, de manera más exacta, $y = \frac{6}{11} + \frac{7}{11}x$, puesto que $5 = \frac{6}{11} + \frac{7}{11}(7)$.

El punto (7, 5) se encuentra sobre la recta $x = -\frac{1}{2} + \frac{3}{2}y$, puesto que $7 = -\frac{1}{2} + \frac{3}{2}(5)$.

Otro método

Las ecuaciones de las dos rectas anteriores son $y = \frac{6}{11} + \frac{7}{11}x$ y $x = -\frac{1}{2} + \frac{3}{2}y$. Si se resuelven de manera simultánea, se encuentra $x = 7$, $y = 5$. Por tanto, las rectas se intersecan en el punto (7, 5).

b) Cuando se hace $x = 12$ en la recta de regresión de *y* en *x*, $y = 0.545 + 0.636(12) = 8.2$.

c) Cuando $y = 3$ en la recta de regresión de *x* en *y*, $x = -0.50 + 1.50(3) = 4.0$.

8.7. Probar que la recta de mínimos cuadrados siempre pasa por el punto $(\bar{x}, \bar{y})$.

Caso 1

x es la variable independiente.

La ecuación de la recta de mínimos cuadrados es (1) $y = a + bx$

La ecuación normal para la recta de mínimos cuadrados es (2) $\sum y = an + b\sum x$

Dividiendo ambos lados de (2) entre *n*, se tiene (3) $\bar{y} = a + b\bar{x}$

Restando (3) de (1), la recta de mínimos cuadrados la escribimos como

(4) $$y - \bar{y} = b(x - \bar{x})$$

lo que indica que la recta pasa por el punto $(\bar{x}, \bar{y})$.

$$- \quad = b(x - \bar{x})$$

lo que indica que la recta pasa por el punto $(\bar{x}, \bar{y})$.

Observemos que, por lo general, las rectas (4) y (5) no son coincidentes, no obstante se intersecan en el punto $(\bar{x}, \bar{y})$.

8.8. Probar que es posible escribir la recta de regresión de mínimos cuadrados de y en x en la forma de (8), en la página 267.

Se tiene de (4) del problema 8.7 que $y - \bar{y} = b(x - \bar{x})$. A partir de la segunda ecuación en (5), página 267,

$$(1) \qquad b = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2}$$

Ahora

$$\begin{aligned}\sum(x - \bar{x})^2 &= \sum(x^2 - 2\bar{x}x + \bar{x}^2)\\ &= \sum x^2 - 2\bar{x}\sum x + \sum \bar{x}^2\\ &= \sum x^2 - 2n\bar{x}^2 + n\bar{x}^2\\ &= \sum x^2 - n\bar{x}^2\\ &= \sum x^2 - \frac{1}{n}\left(\sum x\right)^2\\ &= \frac{1}{n}\left[n\sum x^2 - \left(\sum x\right)^2\right]\end{aligned}$$

También

$$\begin{aligned}\sum(x - \bar{x})(y - \bar{y}) &= \sum(xy - \bar{x}y - \bar{y}x + \bar{x}\bar{y})\\ &= \sum xy - \bar{x}\sum y - \bar{y}\sum x + \sum \bar{x}\bar{y}\\ &= \sum xy - n\bar{x}\bar{y} - n\bar{y}\bar{x} + n\bar{x}\bar{y}\\ &= \sum xy - n\bar{x}\bar{y}\\ &= \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}\\ &= \frac{1}{n}\left[n\sum xy - \left(\sum x\right)\left(\sum y\right)\right]\end{aligned}$$

Por tanto, (1) se convierte en

$$b = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2}$$

a partir de lo cual se obtiene el resultado (8). Se prueba (12), página 268, cuando se intercambian x y y.

8.9. Sean $x = x' + h$, $y = y' + k$, donde h y k son constantes cualesquiera. Demostrar que

$$b = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} = \frac{n\sum x'y' - \left(\sum x'\right)\left(\sum y'\right)}{n\sum x'^2 - \left(\sum x'\right)^2}$$

Del problema 8.8, se tiene

$$b = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2}$$

$= x' + h, y = y' + k$, tenemos

$$\bar{x} = \bar{x}' + h, \quad \bar{y} = \bar{x}' + k$$

En consecuencia,

$$\frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2} = \frac{\sum(x' - \overline{x'})(y' - \overline{y'})}{\sum(x' - \overline{x'})^2}$$

$$= \frac{n\sum x'y' - \left(\sum x'\right)\left(\sum y'\right)}{n\sum x'^2 - \left(\sum x'\right)^2}$$

El resultado es útil en el desarrollo de un atajo para obtener rectas de mínimos cuadrados al restar constantes adecuadas de los valores dados de x y y (vea el problema 8.12).

8.10. Si, en particular, $h = \bar{x}$, $k = \bar{y}$ en el problema 8.9, demostrar que

$$b = \frac{\sum x'y'}{\sum x'^2}$$

Lo anterior se deduce de inmediato del problema 8.9, puesto que

$$\sum x' = \sum(x - \bar{x}) = \sum x - n\bar{x} = 0$$

y de manera similar, $\sum y' = 0$.

8.11. La tabla 8-3 muestra las estaturas respectivas x y y de una muestra de 12 padres y sus hijos mayores. *a*) Construir un diagrama de dispersión. *b*) Determinar la recta de regresión de mínimos cuadrados de y en x. *c*) Dibujar la recta de regresión de mínimos cuadrados de x en y.

Tabla 8-3

Estatura de padres (pulgadas) x	65	63	67	64	68	62	70	66	68	67	69	71
Estatura de hijos (pulgadas) y	68	66	68	65	69	66	68	65	71	67	68	70

a) Se obtiene el diagrama de dispersión cuando se grafican los puntos (x, y) en un sistema de coordenadas rectangulares como se muestra en la figura 8-9.

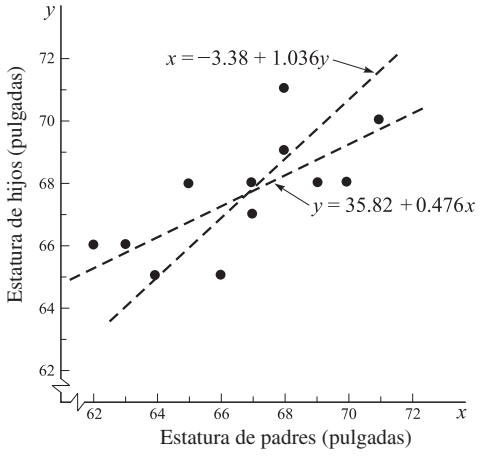


Figura 8-9

$= a + bx$, donde se obtienen a y b al resolver las ecuaciones normales

$$\sum y = an + b\sum x$$
$$\sum xy = a\sum x + b\sum x^2$$

Se muestran las sumas en la tabla 8-4, por lo que las ecuaciones normales se convierten en

$$12a + 800b = 811$$
$$800a + 53\,418b = 54\,107$$

A partir de lo cual se encuentra $a = 35.82$ y $b = 0.476$, por lo que $y = 35.82 + 0.476x$. La gráfica de esta ecuación se observa en la figura 8-9.

Otro método

$$a = \frac{\left(\sum y\right)\left(\sum x^2\right) - \left(\sum x\right)\left(\sum xy\right)}{n\sum x^2 - \left(\sum x\right)^2} = 35.82, \qquad b = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} = 0.476$$

Tabla 8-4

x	y	x^2	xy	y^2
65	68	4 225	4 420	4 624
63	66	3 969	4 158	4 356
67	68	4 489	4 556	4 624
64	65	4 096	4 160	4 225
68	69	4 624	4 692	4 761
62	66	3 844	4 092	4 356
70	68	4 900	4 760	4 624
66	65	4 356	4 290	4 225
68	71	4 624	4 828	5 041
67	67	4 489	4 489	4 489
69	68	4 761	4 692	4 624
71	70	5 041	4 970	4 900
$\sum x = 800$	$\sum y = 811$	$\sum x^2 = 53\,418$	$\sum = 54\,107$	$\sum y^2 = 54\,849$

c) La recta de regresión de x en y está dada por $x = c + dy$, donde c y d se obtienen resolviendo las ecuaciones normales

$$\sum x = cn + d\sum y$$
$$\sum xy = c\sum y + d\sum y^2$$

Utilizando las sumas de la tabla 8-4, éstas se transforman en

$$12c + 811d = 800$$
$$811c + 54\,849d = 54\,107$$

a partir de lo cual encontramos $c = -3.38$ y $d = 1.036$, por lo que $x = -3.38 + 1.036y$. La gráfica de esta ecuación se muestra en la figura 8-9.

Otro método

$$c = \frac{\left(\sum x\right)\left(\sum y^2\right) - \left(\sum y\right)\left(\sum xy\right)}{n\sum y^2 - \left(\sum y\right)^2} = -3.38, \qquad d = \frac{n\sum xy - \left(\sum y\right)\left(\sum x\right)}{n\sum y^2 - \left(\sum y\right)^2} = 1.036$$

$$= \frac{n\sum x'y' - \left(\sum x'\right)\left(\sum y'\right)}{n\sum x'^2 - \left(\sum x'\right)^2} = \frac{(12)(47) - (-16)(-5)}{(12)(106) - (16)^2} = 0.476$$

Además, puesto que $x' = x - 68$, $y' = y - 68$, tenemos $\bar{x}' = \bar{x} - 68$, $\bar{y}' = \bar{y} - 68$. De tal modo,

$$\bar{x} = \bar{x}' + 68 = -\frac{16}{12} + 68 = 66.67, \qquad \bar{y} = \bar{y}' + 68 = -\frac{5}{12} + 68 = 67.58$$

La ecuación de regresión que buscamos de y en x es $y - \bar{y} = b(x - \bar{x})$. Esto es,

$$y - 67.58 = 0.476(x - 66.07) \quad \text{o} \quad y = 35.85 + 0.476x$$

de acuerdo con el problema 8.11, excepto por errores de redondeo. De manera similar se obtiene la ecuación de regresión de x en y.

Tabla 8-5

x'	y'	x'^2	$x'y'$	y'^2
−3	0	9	0	0
−5	−2	25	10	4
−1	0	1	0	0
−4	−3	16	12	9
0	1	0	0	1
−6	−2	36	12	4
2	0	4	0	0
−2	−3	4	6	9
0	3	0	0	9
−1	−1	1	1	1
1	0	1	0	0
3	2	9	6	4
$\sum x' = -16$	$\sum y' = -5$	$\sum x'^2 = 106$	$\sum x'y' = 47$	$\sum y'^2 = 41$

ECUACIONES NO LINEALES REDUCIBLES A LA FORMA LINEAL

8.13. La tabla 8-6 proporciona valores experimentales de la presión P de cierta masa de gas que corresponde a varios valores del volumen V. De acuerdo con los principios termodinámicos, existe una relación entre las variables de la forma $PV^\gamma = C$, donde γ y C son constantes. *a*) Determinar los valores de γ y C. *b*) Escriba la ecuación que relaciona P y V. *c*) Estime P cuando $V = 100.0$ pulg^3.

Tabla 8-6

Volumen V (pulg^3)	54.3	61.8	72.4	88.7	118.6	194.0
Presión P (lb/pulg^2)	61.2	49.5	37.6	28.4	19.2	10.1

Puesto que $PV^\gamma = C$, al tomar los logaritmos de base 10, se tiene

$$\log P + \gamma \log V = \log C \qquad \text{o} \qquad \log P = \log C - \gamma \log V$$

$$= a + bx$$

donde $a = \log C$ y $b = -\gamma$.

La tabla 8-7 proporciona los valores de x y y que corresponden a los valores de V y P en la tabla 8-6 y también indica los cálculos necesarios para encontrar la recta de mínimos cuadrados (1).

Tabla 8-7

$x = \log V$	$y = \log P$	x^2	xy
1.7348	1.7868	3.0095	3.0997
1.7910	1.6946	3.2077	3.0350
1.8597	1.5752	3.4585	2.9294
1.9479	1.4533	3.7943	2.8309
2.0741	1.2833	4.3019	2.6617
2.2878	1.0043	5.2340	2.2976
$\sum x = 11.6953$	$\sum y = 8.7975$	$\sum x^2 = 23.0059$	$\sum xy = 16.8543$

Las ecuaciones normales correspondientes a la recta de mínimos cuadrados (1) son

$$\sum y = an + b\sum x \qquad \sum xy = a\sum x + b\sum x^2$$

de las cuales

$$a = \frac{\left(\sum y\right)\left(\sum x^2\right) - \left(\sum x\right)\left(\sum xy\right)}{n\sum x^2 - \left(\sum x\right)^2} = 4.20, \qquad b = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} = -1.40$$

Entonces $y = 4.20 - 1.40x$.

a) Puesto que $a = 4.20 = \log C$ y $b = -1.40 = -\gamma$, $C = 1.60 \times 10^4$ y $\gamma = 1.40$.

b) $PV^{1.40} = 16\,000$.

c) Cuando $V = 100$, $x = \log V = 2$ y $y = \log P = 4.20 - 1.40(2) = 1.40$. Entonces, $P = \text{antilog } 1.40 = 25.1$ libras/pulg^2.

8.14. Resolver el problema 8.13 graficando los datos en papel log-log.

Para cada par de valores de la presión P y de volumen V en la tabla 8-6, se obtiene un punto que se grafica en el *papel log-log* especialmente construido y que se muestra en la figura 8-10.

También se indica una recta (trazada a mano) que aproxima estos puntos. La gráfica resultante muestra que existe una relación lineal entre $\log P$ y $\log V$, lo cual se representa mediante la ecuación

$$\log P = a + b \log V \qquad \text{o} \qquad y = a + bx$$

La pendiente b, que es negativa en este caso, está dada numéricamente por la proporción entre la longitud de AB y la longitud de AC. La medición en este caso produce $b = -1.4$.

Para obtener a, se necesita un punto en la recta. Por ejemplo, cuando $V = 100$, $P = 25$ a partir de la gráfica. Entonces,

$$a = \log P - b \log V = \log 25 + 1.4 \log 100 = 1.4 + (1.4)(2) = 4.2$$

por lo que

$$\log P + 1.4 \log V = 4.2, \qquad \log PV^{1.4} = 4.2, \qquad \text{y} \qquad PV^{1.4} = 16$$

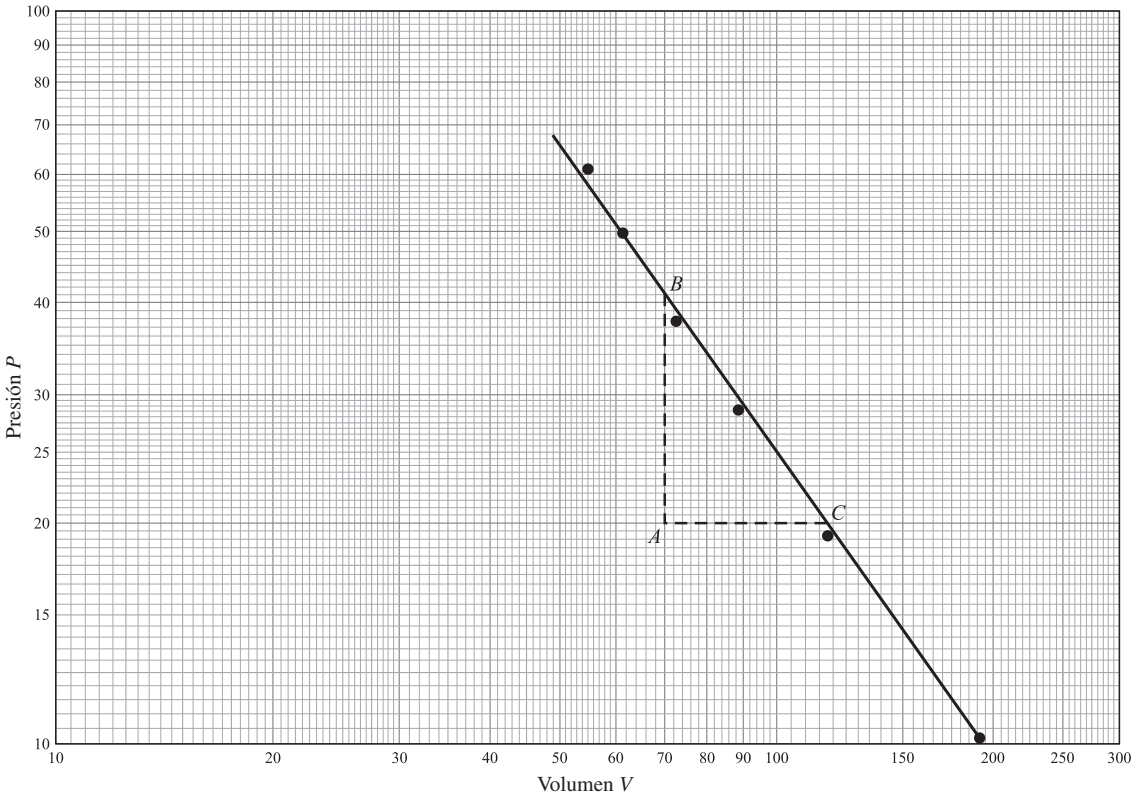


Figura 8-10

PARÁBOLA DE MÍNIMOS CUADRADOS

8.15. Deducir las ecuaciones normales (19), página 269, de la parábola de mínimos cuadrados.

$$y = a + bx + cx^2$$

Sean los puntos muestrales (x_1, y_1), (x_2, y_2), . . . , (x_n, y_n). Entonces, los valores de y en la parábola de mínimos cuadrados correspondientes a $x_1, x_2, \ldots, x_n$ son

$$a + bx_1 + cx_1^2, \qquad a + bx_2 + cx_2^2, \qquad \ldots, \qquad a + bx_n + cx_n^2$$

Por tanto, las desviaciones de $y_1, y_2, \ldots, y_n$ están dadas por

$$d_1 = a + bx_1 + cx_1^2 - y_1, \qquad d_2 = a + bx_2 + cx_2^2 - y_2, \qquad \ldots, \qquad d_n = a + bx_n + cx_n^2 - y_n$$

y la suma de los cuadrados de las desviaciones está dada por

$$\sum d^2 = \sum (a + bx + cx^2 - y)^2$$

Ésta es una función de a, b y c, es decir,

$$F(a, b, c) = \sum (a + bx + cx^2 - y)^2$$

Para minimizar esta función, necesitamos

$$\frac{\partial F}{\partial a} = 0, \qquad \frac{\partial F}{\partial b} = 0, \qquad \frac{\partial F}{\partial c} = 0$$

Ahora,

$$\frac{\partial F}{\partial a} = \sum \frac{\partial}{\partial a}(a + bx + cx^2 - y)^2 = \sum 2(a + bx + cx^2 - y)$$

$$\frac{\partial F}{\partial b} = \sum \frac{\partial}{\partial b}(a + bx + cx^2 - y)^2 = \sum 2x(a + bx + cx^2 - y)$$

$$\frac{\partial F}{\partial c} = \sum \frac{\partial}{\partial c}(a + bx + cx^2 - y)^2 = \sum 2x^2(a + bx + cx^2 - y)$$

Simplificando cada una de estas sumas e igualando a cero, se producen las ecuaciones (19), página 269.

$$\sum y = an + b\sum x + c\sum x^2$$
$$\sum xy = a\sum x + b\sum x^2 + c\sum x^3$$
$$\sum \quad = a\sum x^2 + b\sum x^3 + c\sum x^4$$

Se ordena el trabajo para calcular las sumas como se muestra en la tabla 8-9.

Tabla 8-9

x	y	x^2	x^3	x^4	xy	x^2y
1.2	4.5	1.44	1.73	2.08	5.40	6.48
1.8	5.9	3.24	5.83	10.49	10.62	19.12
3.1	7.0	9.61	29.79	92.35	21.70	67.27
4.9	7.8	24.01	117.65	576.48	38.22	187.28
5.7	7.2	32.49	185.19	1055.58	41.04	233.93
7.1	6.8	50.41	357.91	2541.16	48.28	342.79
8.6	4.5	73.96	636.06	5470.12	38.70	332.82
9.8	2.7	96.04	941.19	9223.66	26.46	259.31
$\sum x =$ 42.2	$\sum y =$ 46.4	$\sum x^2 =$ 291.20	$\sum x^3 =$ 2 275.35	$\sum x^4 =$ 18 971.92	$\sum xy =$ 230.42	$\sum x^2y =$ 1 449.00

Entonces, puesto que $n = 8$, las ecuaciones normales (1) se convierten en

$$8a + 42.2b + 291.20c = 46.4$$

(2)
$$42.2a + 291.20b + 2\,275.35c = 230.42$$
$$291.20a + 2\,275.35b + 18\,971.92c = 1\,449.00$$

Resolviendo, $a = 2.588$, $b = 2.065$, $c = -0.2110$, por lo que la parábola de mínimos cuadrados tiene la ecuación

$$y = 2.588 + 2.065x - 0.2110x^2$$

8.17. Emplear la parábola de mínimos cuadrados del problema 8.16 para estimar los valores de y a partir de los valores dados de x.

Para $x = 1.2$, $y_{est} = 2.588 + 2.065(1.2) - 0.2110(1.2)^2 = 4.762$. De manera similar, se obtienen otros valores estimados. Los resultados se muestran en la tabla 8-10 junto con los valores reales de y.

Tabla 8-10

y_{est}	4.762	5.621	6.962	7.640	7.503	6.613	4.741	2.561
y	4.5	5.9	7.0	7.8	7.2	6.8	4.5	2.7

REGRESIÓN MÚLTIPLE

8.18. Una variable z se calcula a partir de las variables x y y por medio de una ecuación de regresión de la forma $z = a + bx + cy$. Demostrar que se obtiene la ecuación de regresión de mínimos cuadrados determinando a, b y c de modo que se satisfaga (21), página 269.

$$+ bx_1 + cy_1, \qquad \ldots, \qquad a + bx_n + cy_n$$

Por tanto, las desviaciones de $z_1, \ldots, z_n$ están dadas por

$$d_1 = a + bx_1 + cy_1 - z_1, \qquad \ldots, \qquad d_n = a + bx_n + cy_n - z_n$$

y la suma de los cuadrados de las desviaciones está dada por

$$\sum d^2 = \sum (a + bx + cy - z)^2$$

Considerando esto como una función de a, b, c e igualando a cero las derivadas parciales con respecto a a, b y c, obtenemos las ecuaciones normales que se requieren (21) en la página 269.

8.19. La tabla 8-11 muestra los pesos z a la libra más cercana, estaturas x a la pulgada más cercana y las edades y al año más cercano de 12 niños. *a*) Encontrar la ecuación de regresión de mínimos cuadrados de z en x y y. *b*) Determinar los valores estimados de z a partir de los valores de x y y. *c*) Estimar el peso de un niño de 9 años de edad y 54 pulgadas de estatura.

Tabla 8-11

Peso (z)	64	71	53	67	55	58	77	57	56	51	76	68
Altura (x)	57	59	49	62	51	50	55	48	52	42	61	57
Edad (y)	8	10	6	11	8	7	10	9	10	6	12	9

a) Se escribe la ecuación de regresión lineal de z en x y y como

$$z = a + bx + cy$$

Las ecuaciones normales (21), página 269, están dadas por

$$\begin{aligned} \sum z &= na + b\sum x + c\sum y \\ \sum xz &= a\sum x + b\sum x^2 + c\sum xy \\ \sum yz &= a\sum y + b\sum xy + c\sum y^2 \end{aligned} \tag{1}$$

El trabajo de calcular las sumas se arregla como se muestra en la tabla 8-12.

Tabla 8-12

z	x	y	z^2	x^2	y^2	xz	yx	xy
64	57	8	4 096	3 249	64	3 648	512	456
71	59	10	5 041	3 481	100	4 189	710	590
53	49	6	2 809	2 401	36	2 597	318	294
67	62	11	4 489	3 844	121	4 154	737	682
55	51	8	3 025	2 601	64	2 805	440	408
58	50	7	3 364	2 500	49	2 900	406	350
77	55	10	5 929	3 025	100	4 235	770	550
57	48	9	3 249	2 304	81	2 736	513	432
56	52	10	3 136	2 704	100	2 912	560	520
51	42	6	2 601	1 764	36	2 142	306	252
76	61	12	5 776	3 721	144	4 636	912	732
68	57	9	4 624	3 249	81	3 876	612	513
$\sum z =$ 753	$\sum x =$ 643	$\sum y =$ 106	$\sum z^2 =$ 48 139	$\sum x^2 =$ 34 843	$\sum y^2 =$ 976	$\sum xz =$ 40 830	$\sum yz =$ 6 796	$\sum xy =$ 5 779

$$12a + 643b + 106c = 753$$
$$643a + 34\,843b + 5\,779c = 40\,830$$
$$+ 5\,779b + 976c = 6\,796$$

Resolviendo $a = 3.6512$, $b = 0.8546$, $c = 1.5063$, y la ecuación de regresión es

(3) $$z = 3.65 + 0.855x + 1.506y$$

b) Usando la ecuación de regresión (3), obtenemos los valores estimados de z, denotados por z_{est}, al sustituir los valores correspondientes de x y y. Los resultados se presentan en la tabla 8-13 junto con los valores muestrales de z.

Tabla 8-13

z_{est}	64.414	69.136	54.564	73.206	59.286	56.925	65.717	58.229	63.153	48.582	73.857	65.920
z	64	71	53	67	55	58	77	57	56	51	76	68

c) Sea $x = 54$ y $y = 9$ en (3), el peso estimado es $z_{est} = 63.356$, o aproximadamente 63 libras.

ERROR ESTÁNDAR DE LA ESTIMACIÓN

8.20. Si la recta de regresión de mínimos cuadrados de y en x está dada por $y = a + bx$, demostrar que el error estándar de estimación, $s_{y.x}$, está dado por

$$s_{y.x}^2 = \frac{\sum y^2 - a\sum y - b\sum xy}{n}$$

Los valores de y según se estima de la recta de regresión están dados por $y_{est} = a + bx$. Entonces,

$$s_{y.x}^2 = \frac{\sum(y - y_{est})^2}{n} = \frac{\sum(y - a - bx)^2}{n}$$
$$= \frac{\sum y(y - a - bx) - a\sum(y - a - bx) - b\sum x(y - a - bx)}{n}$$

Pero
$$\sum(y - a - bx) = \sum y - an - b\sum x = 0$$
$$\sum x(y - a - bx) = \sum xy - a\sum x - b\sum x^2 = 0$$

puesto que a partir de las ecuaciones normales

$$\sum y = an + b\sum x \qquad \sum xy = a\sum x + b\sum x^2$$

Entonces,
$$s_{y.x}^2 = \frac{\sum y(y - a - bx)}{n} = \frac{\sum y^2 - a\sum y - b\sum xy}{n}$$

Es posible extender este resultado a ecuaciones de regresión no lineales.

8.21. Demostrar que es posible escribir el resultado del problema 8.20 como

$$s_{y.x}^2 = \frac{\sum(y - \bar{y})^2 - b\sum(x - \bar{x})(y - \bar{y})}{n}$$

Método 1

Sea $x = x' + \bar{x}$, $y = y' + \bar{y}$. Entonces, del problema 8.20,

$$ns_{y.x}^2 = \sum y^2 - a\sum y - b\sum xy$$
$$= \sum(y' + \bar{y})^2 - a\sum(y' + \bar{y}) - b\sum(x' + \bar{x})(y' + \bar{y})$$
$$= \sum(y'^2 + 2y'\bar{y} + \bar{y}^2) - a\left(\sum y' + n\bar{y}\right) - b\sum(x'y' + \bar{x}y' + x'\bar{y} + \bar{x}\bar{y})$$

$$= \sum y'^2 + 2\bar{y}\sum y' + n\bar{y}^2 - an\bar{y} - b\sum x'y' - b\bar{x}\sum y' - b\bar{y}\sum x' - bn\bar{x}\bar{y}$$
$$= \sum y'^2 + n\bar{y}^2 - an\bar{y} - b\sum x'y' - bn\bar{x}\bar{y}$$
$$= \sum y'^2 - b\sum x'y' + n\bar{y}(\bar{y} - a - b\bar{x})$$
$$= \sum y'^2 - b\sum x'y'$$
$$= \sum (y - \bar{y})^2 - b\sum (x - \bar{x})(y - \bar{y})$$

donde hemos usado los resultados $\sum x' = 0$, $\sum y' = 0$ y $\bar{y} = a + b\bar{x}$ (que se deducen cuando se dividen los dos lados de la ecuación normal $\sum y = an + b\sum x$ entre n), lo que demuestra el resultado.

Método 2

Sabemos que la recta de regresión se puede escribir como $y - \bar{y} = b(x - \bar{x})$, lo que corresponde a empezar con $y = a + bx$ y después sustituir a por cero, x por $x - \bar{x}$ y y por $y - \bar{y}$. Cuando se efectúan estas sustituciones en el problema 8.20, se obtiene el resultado que se buscaba.

8.22. Calcular el error estándar de estimación, $s_{y.x}$, de los datos del problema 8.11.

A partir del problema 8.11*b*), la recta de regresión de y en x es $y = 35.82 + 0.476x$. La tabla 8-14 lista los valores verdaderos de y (de la tabla 8-3) y los valores estimados de y, denotados por y_{est}, como se obtuvieron de la recta de regresión. Por ejemplo, si $x = 65$, obtenemos $y_{est} = 35.82 + 0.476(65) = 66.76$.

Tabla 8-14

x	65	63	67	64	68	62	70	66	68	67	69	71
y	68	66	68	65	69	66	68	65	71	67	68	70
y_{est}	66.76	65.81	67.71	66.28	68.19	65.33	69.14	67.24	68.19	67.71	68.66	69.62
$y - y_{est}$	1.24	0.19	0.29	−1.28	0.81	0.67	−1.14	−2.24	2.81	−0.71	−0.66	0.38

También se listan los valores $y - y_{est}$, que son necesarios para calcular $s_{y.x}$,

$$s_{y.x}^2 = \frac{\sum (y - y_{est})^2}{n} = \frac{(1.24)^2 + (0.19) + \cdots + (0.38)^2}{12} = 1.642$$

y $s_{y.x} = \sqrt{1.642} = 1.28$ pulgadas.

8.23. *a*) Construir dos rectas paralelas a la recta de regresión del problema 8.11 y a una distancia vertical de $s_{y.x}$ de ella. *b*) Determinar el porcentaje de puntos (datos) entre estas dos rectas.

a) En la figura 8-11 se muestra como una recta sólida la recta de regresión $y = 35.82 + 0.476x$ de acuerdo a como se obtuvo en el problema 8.11. La misma figura presenta, como rectas punteadas, las dos rectas paralelas, cada una con una distancia vertical a la recta de regresión de $s_{y.x} = 1.28$ (vea el problema 8.22).

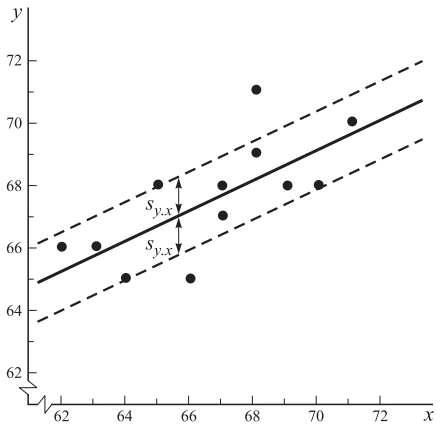


Figura 8-11

$$= \sqrt{n/(n-2)}s_{y.x} = \sqrt{12/10}(1.28) = 1.40$$

COEFICIENTE DE CORRELACIÓN LINEAL

8.24. Demostrar que $\sum(y - \bar{y})^2 = \sum(y - y_{est})^2 + \sum(y_{est} - \bar{y})^2$.

Si se elevan al cuadrado los dos lados de $y - \bar{y} = (y - y_{est}) + (y_{est} - \bar{y})$ y después se suman, se tiene

$$\sum(y - \bar{y})^2 = \sum(y - y_{est})^2 + \sum(y_{est} - \bar{y})^2 + 2\sum(y - y_{est})(y_{est} - \bar{y})$$

El resultado que buscamos se deduce de inmediato si se demuestra que la última suma es cero. Lo anterior es así en el caso de la regresión lineal, puesto que

$$\begin{aligned}\sum(y - y_{est})(y_{est} - \bar{y}) &= \sum(y - a - bx)(a + bx - \bar{y}) \\ &= a\sum(y - a - bx) + b\sum x(y - a - bx) - \bar{y}\sum(y - a - bx) \\ &= 0\end{aligned}$$

en virtud de las ecuaciones normales $\sum(y - a - bx) = 0$, $\sum x(y - a - bx) = 0$.

De manera similar, la validez del resultado puede demostrarse para la regresión no lineal utilizando una curva de mínimos cuadrados dada por $y_{est} = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$.

8.25. Calcular *a*) la variación explicada, *b*) la variación no explicada, *c*) la variación total para los datos del problema 8.11.

Se tiene $\bar{y} = 67.58$ del problema 8.12 (o de la tabla 8-4, ya que $\bar{y} = 811/12 = 67.58$). Empleando los valores y_{est} de la tabla 8-14, se construye la tabla 8-15.

Tabla 8-15

$y_{est} - \bar{y}$	−0.82	−1.77	0.13	−1.30	0.61	−2.25	1.56	−0.34	0.61	0.13	1.08	2.04

a) Variación explicada = $\sum(y_{est} - \bar{y})^2 = (-0.82)^2 + \cdots + (2.04)^2 = 19.22$.

b) Variación no explicada = $\sum(y - y_{est})^2 = ns_{y.x}^2 = 19.70$, del problema 8.22.

c) Variación total = $\sum(y - \bar{y})^2 = 19.22 + 19.70 = 38.92$, del problema 8.24.
Los resultados en *b*) y *c*) también se obtienen por cálculo directo de la suma de cuadrados.

8.26. Para los datos del problema 8.11, calcular *a*) el coeficiente de determinación y *b*) el coeficiente de correlación. Utilizar los resultados del problema 8.25.

a) Coeficiente de determinación = $r^2 = \dfrac{\text{variación explicada}}{\text{variación total}} = \dfrac{19.22}{38.92} = 0.4938$.

b) Coeficiente de correlación = $r = \pm\sqrt{0.4938} = \pm 0.7027$.

Puesto que la variable y_{est} aumenta cuando aumenta x, la correlación es positiva, por lo que se escribe $r = 0.7027$ o sea 0.70 con dos cifras significativas.

$y_{est} = a + bx = bx'$, donde $b = \sum x'y'/\sum x'^2$, $x' = x - \bar{x}$ y $y'_{est} = y_{est} - \bar{y}$. Entonces, usando $y' = y - \bar{y}$, tenemos

$$r^2 = \frac{\text{variación explicada}}{\text{variación total}} = \frac{\sum(y_{est} - \bar{y})^2}{\sum(y - \bar{y})^2} = \frac{\sum y'^2_{est}}{\sum y'^2}$$

$$= \frac{\sum b^2x'^2}{\sum y'^2} = \frac{b^2\sum x'^2}{\sum y'^2} = \left(\frac{\sum x'y'}{\sum x'^2}\right)^2\left(\frac{\sum x'^2}{\sum y'^2}\right) = \frac{\left(\sum x'y'\right)^2}{\sum x'^2\sum y'^2}$$

por lo que

$$r = \pm\frac{\sum x'y'}{\sqrt{\sum x'^2\sum y'^2}}$$

Sin embargo, puesto que $\sum x'y'$ es positiva cuando y_{est} aumenta a medida que crece x, pero negativa cuando y_{est} disminuye a medida que aumenta x, la expresión de r automáticamente toma el signo correcto. En consecuencia, se deduce el resultado que buscamos.

8.28. Con base en la fórmula momento-producto, obtener el coeficiente de correlación lineal para los datos del problema 8.11.

Se organiza el trabajo para calcular lo anterior como se muestra en la tabla 8-16. Entonces,

$$r = \frac{\sum x'y'}{\sqrt{\left(\sum x'^2\right)\left(\sum y'^2\right)}} = \frac{40.34}{\sqrt{(84.68)(38.92)}} = 0.7027$$

que concuerda con el problema 8.26*b*).

Tabla 8-16

x	y	$x' = x - \bar{x}$	$y' = y - \bar{y}$	x'^2	$x'y'$	y'^2
65	68	−1.7	0.4	2.89	−0.68	0.16
63	66	−3.7	−1.6	13.69	5.92	2.56
67	68	0.3	0.4	0.09	0.12	0.16
64	65	−2.7	−2.6	7.29	7.02	6.76
68	69	1.3	1.4	1.69	1.82	1.96
62	66	−4.7	−1.6	22.09	7.52	2.56
70	68	3.3	0.4	10.89	1.32	0.16
66	65	−0.07	−2.6	0.49	1.82	6.76
68	71	1.3	3.4	1.69	4.42	11.56
67	67	0.3	−0.6	0.09	−0.18	0.36
69	68	2.3	0.4	5.29	0.92	0.16
71	70	4.3	2.4	18.49	10.32	5.76
$\sum x = 800$ $\bar{x} = 800/12$ $= 66.7$	$\sum y = 811$ $\bar{y} = 811/12$ $= 67.6$			$\sum x'^2 =$ 84.68	$\sum x'y' =$ 40.34	$\sum y'^2 =$ 38.92

$$= a + bx \quad \text{donde} \quad b = \frac{rs_y}{s_x}$$

De manera similar, la recta de regresión de x en y es

$$x = c + dy \quad \text{donde} \quad d = \frac{rs_x}{s_y}$$

Entonces,
$$bd = \left(\frac{rs_y}{s_x}\right)\left(\frac{rs_x}{s_y}\right) = r^2$$

8.30. Usar el resultado del problema 8.29 para calcular el coeficiente de correlación correspondiente a los datos del problema 8.11.

De los problemas 8.11*b*) y 8.11*c*), respectivamente,

$$b = \frac{484}{1\,016} = 0.476 \qquad d = \frac{484}{467} = 1.036$$

Entonces,
$$r^2 = bd = \left(\frac{484}{1\,016}\right)\left(\frac{484}{467}\right) \quad \text{o} \quad r = 0.7027$$

lo que concuerda con los problemas 8.26*b*) y 8.28.

8.31. Demostrar que el coeficiente de correlación lineal está dado por

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

En el problema 8.27 se demostró que

(1)
$$r = \frac{\sum x'y'}{\sqrt{\left(\sum x'^2\right)\left(\sum y'^2\right)}} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\left[\sum (x - \bar{x})^2\right]\left[\sum (y - \bar{y})^2\right]}}$$

Pero
$$\sum (x - \bar{x})(y - \bar{y}) = \sum (xy - \bar{x}y - x\bar{y} + \bar{x}\bar{y}) = \sum xy - \bar{x}\sum y - \bar{y}\sum x + n\bar{x}\bar{y}$$
$$= \sum xy - n\bar{x}\bar{y} - n\bar{y}\bar{x} + n\bar{x}\bar{y} = \sum xy - n\bar{x}\bar{y}$$
$$= \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}$$

puesto que $\bar{x} = (\sum x)/n$ y $\bar{y} = (\sum y)/n$.

De modo similar,
$$\sum (x - \bar{x})^2 = \sum (x^2 - 2x\bar{x} + \bar{x}^2) = \sum x^2 - 2\bar{x}\sum x + n\bar{x}^2$$
$$= \sum x^2 - \frac{2\left(\sum x\right)^2}{n} + \frac{\left(\sum x\right)^2}{n} = \sum x^2 - \frac{\left(\sum x\right)^2}{n}$$

y
$$\sum (y - \bar{y})^2 = \sum y^2 - \frac{\left(\sum y\right)^2}{n}$$

Entonces, (1) se convierte en

$$r = \frac{\sum xy - \left(\sum x\right)\left(\sum y\right)/n}{\sqrt{\left[\sum x^2 - \left(\sum x\right)^2/n\right]\left[\sum y^2 - \left(\sum y\right)^2/n\right]}} = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

$$= \frac{(12)(54 \quad -}{\sqrt{[(12)(53 \quad - \quad -}} = 0.7027$$

como en los problemas 8.26*b*), 8.28 y 8.30.

COEFICIENTE DE CORRELACIÓN GENERALIZADO

8.33. *a*) Calcular el coeficiente de correlación lineal entre las variables x y y del problema 8.16. *b*) Determinar el coeficiente de correlación lineal entre estas variables, suponiendo la relación parabólica que se obtuvo en el problema 8.16. *c*) Explicar la diferencia entre los coeficientes de correlación obtenidos en *a*) y *b*). *d*) ¿Qué porcentaje de la variación total queda sin explicarse por aceptar la suposición de una relación parabólica entre x y y?

a) Con base en los cálculos de la tabla 8-9 y el hecho adicional de que $\sum y^2 = 290.52$, encontramos

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

$$= \frac{(8)(230.42) - (42.2)(46.4)}{\sqrt{[(8)(291.20) - (42.2)^2][(8)(290.52) - (46.4)^2]}} = -0.3743$$

b) A partir de la tabla 8-9, $\bar{y} = (\sum y)/n = (46.4)/8 = 5.80$. Entonces,

$$\text{Variación total} = \sum(y - \bar{y})^2 = 21.40$$

De la tabla 8-10,

$$\text{Variación explicada} = \sum(y_{\text{est}} - \bar{y})^2 = 21.02$$

En consecuencia, $r^2 = \dfrac{\text{variación explicada}}{\text{variación total}} = \dfrac{21.02}{21.40} = 0.9822$ y $r = 0.9911$

c) El hecho de que la parte *a*) muestra un coeficiente de correlación lineal de sólo -0.3743 indica prácticamente que no existe *relación lineal* entre x y y. Sin embargo, existe una muy buena *relación no lineal* que proporciona la parábola del problema 8.16, como se indica a partir del hecho de que el coeficiente de correlación en *b*) es casi 1.

d) $$\frac{\text{Variación no explicada}}{\text{Variación total}} = 1 - r^2 = 1 - 0.9822 = 0.0178$$

En consecuencia, 1.78% de la variación total queda sin explicación. Es posible que ello se deba a fluctuaciones aleatorias o a una variable adicional que no ha sido considerada.

8.34. Calcular *a*) s_y y *b*) $s_{y.x}$ para los datos del problema 8.16.

a) Del problema 8.33*b*), $\sum(y - \bar{y})^2 = 21.40$. Entonces, la desviación estándar de y es

$$s_y = \sqrt{\frac{\sum(y - \bar{y})^2}{n}} = \sqrt{\frac{21.40}{8}} = 1.636 \text{ o sea } 1.64$$

b) **Primer método**

Usando *a*) y el problema 8.33*b*), el error estándar de la estimación de y en x es

$$s_{y.x} = s_y\sqrt{1 - r^2} = 1.636\sqrt{1 - (0.9911)^2} = 0.218 \text{ o sea } 0.22$$

$$= \sqrt{\frac{\sum(y - y_{est})^2}{n}} = \sqrt{\frac{\text{variación no explicada}}{n}} = \sqrt{\frac{21.40 - 21.02}{8}} = 0.218 \text{ o sea } 0.22$$

Tercer método

Usando el problema 8.16 y el cálculo adicional, $\sum y^2 = 290.52$, tenemos

$$s_{y.x} = \sqrt{\frac{\sum y^2 - a\sum y - b\sum xy - c\sum x^2 y}{n}} = 0.218 \text{ o sea } 0.22$$

8.35. Explicar cómo se determina un coeficiente de correlación múltiple para las variables del problema 8.19.

Puesto que z se determina de x y y, es de interés el *coeficiente de correlación múltiple de z en x y y.* Para obtener éste, vea del problema 8.19 que

$$\text{Variación no explicada} = \sum(z - z_{est})^2$$
$$= (64 - 64.414)^2 + \cdots + (68 - 65.920)^2 = 258.88$$

$$\text{Variación total} = \sum(z - \bar{z})^2 = \sum z^2 - n\bar{z}^2$$
$$= 48 \quad - \quad = 888.25$$

$$\text{Variación explicada} = 888.25 - 258.88 = 629.37$$

Entonces,

$$\text{Coeficiente de correlación múltiple de } z \text{ en } x \text{ y } y$$
$$= \sqrt{\frac{\text{variación explicada}}{\text{variación total}}} = \sqrt{\frac{629.37}{888.25}} = 0.8418$$

Debe mencionarse que si se considerara la regresión de x en y y z, el coeficiente de correlación múltiple de x en y y z sería por lo general diferente al valor anterior.

CORRELACIÓN DE RANGOS

8.36. Deducir la fórmula de correlación de rangos de Spearman (36), página 271.

Aquí se consideran los valores nx (esto es, pesos) y de n valores correspondientes de y (esto es, estaturas). Sea x_j el rango dado al valor j-ésimo de x, y y_j el rango dado al valor j-ésimo de y. Los rangos son los enteros 1 a n. La media de x_j es, entonces,

$$\bar{x} = \frac{1 + 2 + \cdots + n}{n} = \frac{n(n + 1)/2}{n} = \frac{n + 1}{2}$$

mientras que la varianza es

$$s_x^2 = \bar{x}^2 - \bar{x}^2 = \frac{1^2 + 2^2 + \cdots + n^2}{n} - \left(\frac{n + 1}{2}\right)^2$$
$$= \frac{n(n + 1)(2n + 1)/6}{n} - \left(\frac{n + 1}{2}\right)^2$$
$$= \frac{n^2 - 1}{12}$$

usando los resultados 1 y 2 del apéndice A. De manera similar, la media $\bar{y}$ y la varianza s_y^2 son iguales a $(n + 1)/2$ y $(n^2 - 1)/12$, respectivamente.

Ahora, si $d_j = x_j - y_j$ son las desviaciones entre los rangos, la varianza de las desviaciones s_d^2 está dada en términos de s_x^2, s_y^2 y el coeficiente de correlación entre los rangos por

$$s_d^2 = s_x^2 + s_y^2 - 2r_{\text{rango}}\, s_x s_y$$

$$= \frac{s_x^2 + s_y^2 - s_d^2}{2s_x s_y}$$

Puesto que $\bar{d} = 0, s_d^2 = (\sum d^2)/n$ y (1) se convierte

$$(2) \qquad r_{\text{rango}} = \frac{(n^2 - 1)/12 + (n^2 - 1)/12 - \left(\sum d^2\right)/n}{(n^2 - 1)/6} = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

8.37. La tabla 8-17 muestra cómo se clasificaron 10 estudiantes de acuerdo con sus logros tanto en el laboratorio como en las clases de un curso de biología. Calcular el coeficiente de la correlación de rangos.

Tabla 8-17

Laboratorio	8	3	9	2	7	10	4	6	1	5
Curso	9	5	10	1	8	7	3	4	2	6

La tabla 8-18 indica las diferencias de rangos d en laboratorio y clase de cada estudiante. Además de proporcionar d^2 y $\sum d^2$.

Tabla 8-18

Diferencia de rangos (d)	−1	−2	−1	1	−1	3	1	2	−1	−1	
d^2	1	4	1	1	1	9	1	4	1	1	$\sum d^2 = 24$

Entonces, $$r_{\text{rango}} = 1 - \frac{6\sum d^2}{n(n^2 - 1)} = 1 - \frac{6(24)}{10(10^2 - 1)} = 0.8545$$

que indica que existe una relación notable entre logros en el laboratorio y en la clase.

8.38. Calcular el coeficiente de la correlación de rangos para los datos del problema 8.11 y comparar el resultado con el coeficiente de correlación que se obtuvo mediante otros métodos.

Las estaturas de los padres en orden ascendente son

$$(1) \qquad 62, 63, 64, 65, 66, 67, 67, 68, 68, 69, 70, 71$$

Puesto que los lugares sexto y séptimo en este arreglo representan la misma estatura (67 pulgadas), se asigna un *rango medio* de 6.5 a los dos lugares. De manera similar, a los lugares octavo y noveno se les asignan el rango de 8.5. Por tanto, a las estaturas de los padres se asignan los rangos

$$(2) \qquad 1, 2, 3, 4, 5, 6.5, 6.5, 8.5, 8.5, 10, 11, 12$$

De igual modo, las estaturas de los hijos ordenadas en orden de magnitud ascendente son

$$(3) \qquad 65, 65, 66, 66, 67, 68, 68, 68, 68, 69, 70, 71$$

y puesto que los lugares sexto, séptimo, octavo y noveno representan la misma estatura (68 pulgadas), se les asigna un *rango medio* de 7.5 $(6 + 7 + 8 + 9)/4$. En consecuencia, se asignan a las estaturas de los hijos los rangos

$$(4) \qquad 1.5, 1.5, 3.5, 3.5, 5, 7.5, 7.5, 7.5, 7.5, 10, 11, 12$$

Empleando las correspondencias (1) y (2), (3) y (4), la tabla 8-3 se convierte en la 8-19.

Tabla 8-19

Rango de padres	4	2	6.5	3	8.5	1	11	5	8.5	6.5	10	12
Rango de hijos	7.5	3.5	7.5	1.5	10	3.5	7.5	1.5	12	5	7.5	11

	– – – – – –	
		$\sum d^2 = 72.50$

Entonces, $$r_{\text{rango}} = 1 - \frac{6\sum d^2}{n(n^2 - 1)} = 1 - \frac{6(72.50)}{12(12^2 - 1)} = 0.7465$$

que concuerda bien con el valor $r = 0.7027$ que se obtuvo en el problema 8.26*b*).

INTERPRETACIÓN PROBABILÍSTICA DE REGRESIÓN Y CORRELACIÓN

8.39. Deducir (39) de (37).

Suponga que la ecuación de regresión es

$$y = E(Y|X = x) = \alpha + \beta x$$

Para el caso de la recta de regresión de mínimos cuadrados es necesario considerar

$$\begin{aligned} E\{[Y - (\alpha + \beta X)]^2\} &= E\{[(Y - \mu_Y) - \beta(X - \mu_X) + (\mu_Y - \beta\mu_X - \alpha)]^2\} \\ &= E[(Y - \mu_Y)^2] + \beta^2 E[(X - \mu_X)^2] - 2\beta E[(X - \mu_X)(Y - \mu_Y)] + (\mu_Y - \beta\mu_X - \alpha)^2 \\ &= \sigma_Y^2 + \beta^2\sigma_X^2 - 2\beta\sigma_{XY} + (\mu_Y - \beta\mu_X - \alpha)^2 \end{aligned}$$

donde hemos usado $E(X - \mu_X) = 0, E(Y - \mu_Y) = 0$.

Denotando la última expresión por $F(\alpha, \beta)$, tenemos

$$\frac{\partial F}{\partial \alpha} = -2(\mu_Y - \beta\mu_X - \alpha), \qquad \frac{\partial F}{\partial \beta} = 2\beta\sigma_X^2 - 2\sigma_{XY} - 2\mu_X(\mu_Y - \beta\mu_X - \alpha)$$

Igualando a cero, que es una condición necesaria para que $F(\alpha, \beta)$ sea un mínimo, se encuentra

$$\mu_Y = \alpha + \beta\mu_X \qquad \beta\sigma_X^2 = \sigma_{XY}$$

Por tanto, si $y = \alpha + \beta x$, entonces $y - \mu_Y = \beta(x - \mu_X)$ o

$$y - \mu_Y = \frac{\sigma_{XY}}{\sigma_X^2}(x - \mu_X)$$

o bien $$\frac{y - \mu_Y}{\sigma_Y} = \rho\left(\frac{x - \mu_X}{\sigma_X}\right)$$

Es importante destacar la similitud de la prueba anterior para poblaciones, en la que se emplean esperanzas, a la prueba correspondiente para muestras, que utiliza sumas. En general, los resultados para muestras tienen resultados análogos en poblaciones y viceversa.

8.40. La función de densidad conjunta de las variables aleatorias X y Y es

$$f(x, y) = \begin{cases} \frac{2}{3}(x + 2y) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{si no es así} \end{cases}$$

Calcular la curva de regresión de mínimos cuadrados de *a*) Y en X, *b*) X en Y.

a) La función de densidad marginal de X es

$$f_1(x) = \int_0^1 \frac{2}{3}(x + 2y)\,dy = \frac{2}{3}(x + 1)$$

$$= \frac{f(x, y)}{f_1(x)} = \begin{cases} \dfrac{x + 2y}{x + 1} & 0 \le y \le 1 \\ 0 & y < 0 \text{ o bien } y > 1 \end{cases}$$

y la curva de regresión de mínimos cuadrados de Y en X está dada por

$$y = E(Y \mid X = x) = \int_{-\infty}^{\infty} y f_2(y \mid x)\, dy$$

$$= \int_0^1 y\left(\frac{x + 2y}{x + 1}\right) dy = \frac{3x + 4}{6x + 6}$$

No se definen $f_2(y \mid x)$ ni la curva de regresión de mínimos cuadrados cuando $x < 0$ o $x > 1$.

b) En el caso de $0 \le y \le 1$, la función de densidad marginal de Y es

$$f_2(y) = \int_0^1 \frac{2}{3}(x + 2y)\, dx = \frac{1}{3}(1 + 4y)$$

Para $0 \le y \le 1$, la densidad condicional de X dada Y es

$$f_1(x \mid y) = \frac{f(x, y)}{f_2(y)} = \begin{cases} \dfrac{2x + 4y}{1 + 4y} & 0 \le x \le 1 \\ 0 & x < 0 \text{ o bien } x > 1 \end{cases}$$

y la curva de regresión de mínimos cuadrados de X en Y es

$$x = E(X \mid Y = y) = \int_{-\infty}^{\infty} x f_1(x \mid y)\, dx$$

$$= \int_0^1 x\left(\frac{2x + 4y}{1 + 4y}\right) dx = \frac{2 + 6y}{3 + 12y}$$

No se define $f_1(x \mid y)$ ni la curva de regresión de mínimos cuadrados cuando $y < 0$ o $y > 1$.

Observemos que las dos curvas de regresión $y = (3x + 4)/(6x + 6)$ y $x = (2 + 6y)/(3 + 12y)$ son diferentes.

8.41. Calcular *a*) $\bar{X}$, *b*) $\bar{Y}$, *c*) σ_X^2, *d*) σ_Y^2, *e*) σ_{XY}, *f*) ρ para la distribución del problema 8.40.

a) $$\bar{X} = \int_{x=0}^{1}\int_{y=0}^{1} x\left[\frac{2}{3}(x + 2y)\right] dx\, dy = \frac{5}{9}$$

b) $$\bar{Y} = \int_{x=0}^{1}\int_{y=0}^{1} y\left[\frac{2}{3}(x + 2y)\right] dx\, dy = \frac{11}{18}$$

c) $$\bar{X}^2 = \int_{x=0}^{1}\int_{y=0}^{1} x^2\left[\frac{2}{3}(x + 2y)\right] dx\, dy = \frac{7}{18}$$

Entonces, $$\sigma_X^2 = \bar{X}^2 - \bar{X}^2 = \frac{7}{18} - \left(\frac{5}{9}\right)^2 = \frac{13}{162}$$

d) $$\bar{Y}^2 = \int_{x=0}^{1}\int_{y=0}^{1} y^2\left[\frac{2}{3}(x + 2y)\right] dx\, dy = \frac{4}{9}$$

Luego $$\sigma_Y^2 = \bar{Y}^2 - \bar{Y}^2 = \frac{4}{9} - \left(\frac{11}{18}\right)^2 = \frac{23}{324}$$

e) $$\bar{X}\bar{Y} = \int_{x=0}^{1}\int_{y=0}^{1} xy\left[\frac{2}{3}(x + 2y)\right] dx\, dy = \frac{1}{3}$$

$$\sigma_{XY} = \overline{XY} - \bar{X}\bar{Y} = \frac{1}{3} - \left(\frac{5}{9}\right)\left(\frac{11}{18}\right) = -\frac{1}{162}$$

f)
$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-1/162}{\sqrt{13/162}\sqrt{23/324}} = -0.0818$$

Observemos que el coeficiente de la correlación lineal es pequeño. Esto es lo que se esperaba de la observación de las rectas de regresión de mínimos cuadrados del problema 8.42.

8.42. Escribir las rectas de regresión de mínimos cuadrados de *a*) Y en X, *b*) X en Y para el problema 8.40.

a) La recta de regresión de Y en X es $\dfrac{y - \bar{Y}}{\sigma_Y} = \rho\left(\dfrac{x - \bar{X}}{\sigma_X}\right)$ o bien

$$y - \bar{Y} = \frac{\sigma_{XY}}{\sigma_X^2}(x - \bar{X}) \quad \text{o} \quad y - \frac{11}{18} = \frac{-1/162}{13/162}\left(x - \frac{5}{9}\right)$$

b) La recta de regresión de X en Y es $\dfrac{x - \bar{X}}{\sigma_X} = \rho\left(\dfrac{y - \bar{Y}}{\sigma_Y}\right)$ o bien

$$x - \bar{X} = \frac{\sigma_{XY}}{\sigma_Y^2}(y - \bar{Y}) \quad \text{o} \quad x - \frac{5}{9} = \frac{-1/162}{23/324}\left(y - \frac{11}{18}\right)$$

TEORÍA MUESTRAL DE LA REGRESIÓN

8.43. En el problema 8.11 encontramos que la ecuación de regresión de y en x era $y = 35.82 + 0.476x$. Demostrar la hipótesis al nivel de significancia de 0.05 de que el coeficiente de regresión de la ecuación de regresión poblacional es tan bajo como 0.180.

$$t = \frac{\beta - b}{s_{y.x}/s_x}\sqrt{n - 2} = \frac{0.476 - 0.180}{1.28/2.66}\sqrt{12 - 2} = 1.95$$

puesto que $s_{y.x} = 1.28$ (cálculo del problema 8.22) y $s_x = \sqrt{\overline{x^2} - (\bar{x})^2} = 2.66$ del problema 8.11.

Con base en una prueba de una cola (unilateral) de la distribución de Student a un nivel de 0.05, se debería rechazar la hipótesis de que el coeficiente de regresión es tan bajo como 0.180 si $t > t_{0.95} = 1.81$ para $12 - 2 = 10$ grados de libertad. Por tanto, podemos rechazar la hipótesis.

8.44. Encontrar los límites de confianza de 95% para el coeficiente de regresión del problema 8.43.

$$\beta = b + \frac{t}{\sqrt{n - 2}}\left(\frac{s_{y.x}}{s_x}\right)$$

Entonces, los límites de confianza de 95% para β (que se obtuvieron al poner $t = \pm t_{0.975} = \pm 2.23$ para $12 - 2 = 10$ grados de libertad) están dados por

$$b \pm \frac{2.23}{\sqrt{12 - 2}}\left(\frac{s_{y.x}}{s_x}\right) = 0.476 \pm \frac{2.23}{\sqrt{10}}\left(\frac{1.28}{2.66}\right) = 0.476 \pm 0.340$$

esto es, existe una confianza de 95% de que β está entre 0.136 y 0.816.

8.45. En el problema 8.11, encontrar los límites de confianza de 95% para las estaturas de los hijos cuyos padres tienen estaturas de *a*) 65.0 y *b*) 70.0 pulgadas.

Puesto que $t_{0.975} = 2.23$ para $12 - 2 = 10$ grados de libertad, los límites de confianza de 95% para y_p son

$$y_0 \pm \frac{2.23}{\sqrt{n - 2}} s_{y.x}\sqrt{n + 1 + \frac{n(x_0 - \bar{x})^2}{s_x^2}}$$

donde $y_0 = 35.82 + 0.476x_0$ (problema 8.11), $s_{y.x} = 1.28$, $s_x = 2.66$ (problema 8.43) y $n = 12$.

$$\pm \frac{2.23}{\sqrt{10}}(1.28)\sqrt{12 + 1 + \frac{12(2.78)}{(2.66)^2}} = 66.76 \pm 3.80 \text{ pulgadas}$$

esto es, tenemos la confianza de alrededor de 95% de que las estaturas de los hijos están entre 63.0 y 70.6 pulgadas.

b) Si $x_0 = 70.0$, $y_0 = 69$ pulgadas. También $(x_0 - x)^2 = (70.0 - 800/12)^2 = 11.11$. Entonces, se calculan los límites de confianza de 95% como 69.14 ± 5.09 pulgadas, esto es, tenemos una confianza de alrededor de 95% de que las estaturas de los hijos estén entre 64.1 y 74.2 pulgadas.

Observemos que para valores grandes de n, los límites de confianza de 95% están dados de manera aproximada por $y_0 \pm 1.96\, s_{y.x}$ o $y_0 \pm 2s_{y.x}$, siempre que $x_0 - \bar{x}$ no sea demasiado grande. Ello concuerda con los resultados aproximados en la página 269. Los métodos de este problema se mantienen sin importar el tamaño de n o $x_0 - \bar{x}$, esto es, los métodos de muestreo son exactos para el caso de una población normal.

8.46. En el problema 8.11, encontrar los límites de confianza de 95% para las estaturas medias de los hijos cuyos padres tienen estaturas de *a*) 65.0, *b*) 70.0 pulgadas.

Puesto que $t_{0.975} = 2.23$ para 10 grados de libertad, los límites de confianza de 95% para $\bar{y}_p$ son

$$y_0 \pm \frac{2.23}{\sqrt{20}} s_{y.x}\sqrt{1 + \frac{(x_0 - \bar{x})^2}{s_x^2}}$$

donde $y_0 = 35.82 + 0.476x_0$ (problema 8.11), $s_{y.x} = 1.28$ (problema 8.43).

a) Si $x_0 = 65.0$, encontramos [comparar con el problema 8.45*a*)] los límites de confianza de 95% de 66.76 ± 1.07 pulgadas, esto es, es posible tener una confianza de 95% de que la estatura media de todos los hijos cuyos padres tienen estaturas de 65.0 pulgadas se ubicará entre 65.7 y 67.8 pulgadas.

b) Si $x_0 = 70.0$, encontramos [comparar con el problema 8.45*b*)] los límites de confianza de 95% de 69.14 ± 1.45 pulgadas, es decir, es posible tener una confianza de 95% de que la estatura media de todos los hijos cuyos padres tienen estaturas de 70.0 pulgadas se ubicará entre 67.7 y 70.6 pulgadas.

TEORÍA DE MUESTREO DE LA CORRELACIÓN

8.47. Con base en una muestra de tamaño 18, se calculó que el coeficiente de correlación es igual a 0.32. ¿Es posible concluir a un nivel de significancia de *a*) 0.05 y *b*) 0.01 que el coeficiente de correlación poblacional que corresponde es significativamente mayor que cero?

Se desea decidir entre las hipótesis (H_0: $\rho = 0$) y (H_1: $\rho > 0$).

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.32\sqrt{18-2}}{\sqrt{1-(0.32)^2}} = 1.35$$

a) Con base en la prueba de una cola de la distribución de Student a un nivel de 0.05, se rechazaría H_0 si $t > t_{0.95} = 1.75$ para $18 - 2 = 16$ grados de libertad. Por tanto, no es posible rechazar H_0 al nivel de 0.05.

b) Debido a que no podemos rechazar H_0 a un nivel de 0.05, es evidente que no es posible rechazarla a un nivel de 0.01.

8.48. ¿Cuál es el tamaño de muestra mínimo necesario para que se concluya que el coeficiente de correlación de 0.32 es significativamente mayor que cero a un nivel de 0.05?

A un nivel de 0.05, con base en la prueba de una cola de la distribución de Student, el valor mínimo de n debe ser tal que

$$\frac{0.32\sqrt{n-2}}{\sqrt{1-(0.32)^2}} = t_{0.95} \quad \text{para} \quad n - 2 \text{ grados de libertad}$$

Para $n = 26$, $\nu = 24$, $t_{0.95} = 1.71$, $t = 0.32\sqrt{24}/\sqrt{1-(0.32)^2} = 1.65$.

Para $n = 27$, $\nu = 25$, $t_{0.95} = 1.71$, $t = 0.32\sqrt{25}/\sqrt{1-(0.32)^2} = 1.69$.

Para $n = 28$, $\nu = 26$, $t_{0.95} = 1.71$, $t = 0.32\sqrt{26}/\sqrt{1-(0.32)^2} = 1.72$.

Entonces, el tamaño de muestra mínimo es $n = 28$.

$$Z = 1.1513 \log\left(\frac{1 + 0.75}{1 - 0.75}\right) = 0.9730, \qquad \mu_Z = 1.1513 \log\left(\frac{1 + 0.60}{1 - 0.60}\right) = 0.6932,$$

$$\sigma_Z = \frac{1}{\sqrt{n - 3}} = \frac{1}{\sqrt{21}} = 0.2182$$

La variable estandarizada es, entonces,

$$z = \frac{Z - \mu_Z}{\sigma_Z} = \frac{0.9730 - 0.6932}{0.2182} = 1.28$$

A un nivel de significancia de 0.05, usando una prueba de una cola de la distribución normal, se debería rechazar la hipótesis sólo si z fuera mayor que 1.64. Por tanto, no es posible rechazar la hipótesis de que el coeficiente de correlación poblacional es tan pequeño como 0.60.

b) Si $\rho = 0.50$, $\mu_Z = 1.1513 \log 3 = 0.5493$ y $z = (0.9730 - 0.5493)/0.2182 = 1.94$. En consecuencia, se puede rechazar la hipótesis de que el coeficiente de correlación poblacional es tan pequeño como $\rho = 0.50$ a un nivel de significancia de 0.05.

8.50. El coeficiente de correlación entre las calificaciones finales de física y matemáticas de un grupo de 21 estudiantes se calculó igual a 0.80. Encontrar los límites de confianza de 95% para este coeficiente.

Puesto que $r = 0.80$ y $n = 21$, los límites de confianza de 95% de μ_2 están dados por

$$Z \pm 1.96\sigma_Z = 1.1513 \log\left(\frac{1 + r}{1 - r}\right) \pm 1.96\left(\frac{1}{\sqrt{n - 3}}\right) = 1.0986 \pm 0.4620$$

Entonces, μ_Z tiene el intervalo de confianza de 95% de 0.5366 a 1.5606.

Si $$\mu_Z = 1.1513 \log\left(\frac{1 + \rho}{1 - \rho}\right) = 0.5366, \qquad \rho = 0.4904.$$

Si $$\mu_Z = 1.1513 \log\left(\frac{1 + \rho}{1 - \rho}\right) = 1.5606, \qquad \rho = 0.9155.$$

Por tanto, los límites de confianza de 95% para ρ son 0.49 y 0.92.

8.51. Dos coeficientes de correlación que se obtuvieron de muestras de tamaños $n_1 = 28$ y $n_2 = 35$ se calcularon como $r_1 = 0.50$ y $r_2 = 0.30$, respectivamente. ¿Hay una diferencia significativa entre los dos coeficientes a un nivel de 0.05?

$$Z_1 = 1.1513 \log\left(\frac{1 + r_1}{1 - r_1}\right) = 0.5493, \qquad Z_2 = 1.1513 \log\left(\frac{1 + r_2}{1 - r_2}\right) = 0.3095$$

y $$\sigma_{Z_1 - Z_2} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} = 0.2669$$

Se desea decidir entre las hipótesis (H_0:$\mu_{Z_1} = \mu_{Z_2}$) y (H_1:$\mu_{Z_1} \neq \mu_{Z_2}$). Con base en la hipótesis H_0,

$$z = \frac{Z_1 - Z_2 - (\mu_{Z_1} - \mu_{Z_2})}{\sigma_{Z_1 - Z_2}} = \frac{0.5493 - 0.3095 - 0}{0.2669} = 0.8985$$

Usando una prueba de dos colas de la distribución normal, se debería rechazar H_0 sólo si $z > 1.96$ o bien $z < -1.96$. Por tanto, no se puede rechazar H_0 y se concluye que los resultados no son significativamente diferentes para un nivel de 0.05.

PROBLEMAS DIVERSOS

8.52. Demostrar la fórmula (25), página 269.

Para el caso de la recta de mínimos cuadrados se tiene, de los problemas 8.20 y 8.21,

$$s_{y.x}^2 = \frac{\sum(y - \bar{y})^2}{n} - b\frac{\sum(x - \bar{x})(y - \bar{y})}{n}$$

Pero, por definición,

$$\frac{\sum(y - \bar{y})^2}{n} = s_y^2 \qquad \frac{\sum(x - \bar{x})(y - \bar{y})}{n} = s_{xy}$$

y, por (6) en la página 267,

$$b = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2} = \frac{s_{xy}}{s_x^2}$$

En consecuencia, $$s_{y.x}^2 = s_y^2 - \frac{s_{xy}^2}{s_x^2} = s_y^2\left[1 - \left(\frac{s_{xy}}{s_x s_y}\right)^2\right] = s_y^2(1 - r^2)$$

Una fórmula análoga se cumple para el caso de la población (vea el problema 8.54).

8.53. Demostrar que $E[(Y - \bar{Y})^2] = E[(Y - Y_{\text{est}})^2] + [(Y_{\text{est}} - \bar{Y})^2]$ para los casos de *a*) una recta de mínimos cuadrados, *b*) una parábola de mínimos cuadrados.

Se tiene $$Y - \bar{Y} = (Y - Y_{\text{est}}) + (Y_{\text{est}} - \bar{Y})$$

Entonces, $$(Y - \bar{Y})^2 = (Y - Y_{\text{est}})^2 + (Y_{\text{est}} - \bar{Y})^2 + 2(Y - Y_{\text{est}})(Y_{\text{est}} - \bar{Y})$$

así que $$E[(Y - \bar{Y})^2 = E[(Y - Y_{\text{est}})^2] + E[(Y_{\text{est}} - \bar{Y})^2] + 2E[(Y - Y_{\text{est}})(Y_{\text{est}} - \bar{Y})]$$

Se obtiene el resultado buscado si demostramos que el último término es cero.

a) Para el caso de la regresión lineal, $Y_{\text{est}} = \alpha + \beta X$. De tal modo que,

$$\begin{aligned} E[(Y - Y_{\text{est}})(Y_{\text{est}} - \bar{Y})] &= E[(Y - \alpha - \beta X)(\alpha + \beta X - \bar{Y})] \\ &= (\alpha - \bar{Y})E(Y - \alpha - \beta X) + \beta E(XY - \alpha X - \beta X^2) \\ &= 0 \end{aligned}$$

debido a las ecuaciones normales

$$E(Y - \alpha - \beta X) = 0, \qquad E(XY - \alpha X - \beta X^2) = 0$$

(Comparar con el problema 8.3.)

b) Para el caso de la regresión parabólica, $Y_{\text{est}} = \alpha + \beta X + \gamma X^2$. Entonces,

$$\begin{aligned} E[(Y - Y_{\text{est}})(Y_{\text{est}} - \bar{Y})] &= E[(Y - \alpha - \beta X - \gamma X^2)(\alpha + \beta X + \gamma X^2 - \bar{Y})] \\ &= (\alpha - \bar{Y})E(Y - \alpha - \beta X - \gamma X^2) + \beta E[X(Y - \alpha - \beta X - \gamma X^2)] \\ &\quad + \gamma E[X^2(Y - \alpha - \beta X - \gamma X^2)] \\ &= 0 \end{aligned}$$

debido a las ecuaciones normales

$$E(Y - \alpha - \beta X - \gamma X^2) = 0, \quad E[X(Y - \alpha - \beta X - \gamma X^2)] = 0, \quad E[X^2(Y - \alpha - \beta X - \gamma X^2)] = 0$$

Comparar con las ecuaciones (19), página 269.

Es posible extender el resultado a curvas de mínimos cuadrados de orden superior.

$$\sigma \quad = \sigma_Y^2(1 - \rho^2) \text{ para la regresión de mínimos cuadrados.}$$

Según la definición del coeficiente de correlación generalizada ρ, junto con el problema 8.53, se tiene, para los casos lineal y parabólico, que

$$\rho^2 = \frac{E[(Y_{\text{est}} - \bar{Y})^2]}{E[(Y - \bar{Y})^2]} = 1 - \frac{E[(Y - Y_{\text{est}})^2]}{E[(Y - \bar{Y})^2]} = 1 - \frac{\sigma_{Y.X}^2}{\sigma_Y^2}$$

y el resultado es inmediato.

La relación se cumple también para el caso de las curvas de mínimos cuadrados de orden superior.

8.55. Demostrar que en el caso de regresión lineal el coeficiente de correlación, según se define mediante (45), se reduce a la definición dada por (40).

El cuadrado del coeficiente de correlación, esto es, el coeficiente *de determinación*, como está dado por (45), en el caso de la regresión lineal corresponde a

$$\rho^2 = \frac{E[(Y_{\text{est}} - \bar{Y})^2]}{E[(Y - \bar{Y})^2]} = \frac{E[(\alpha + \beta X - \bar{Y})^2]}{\sigma_Y^2} \tag{1}$$

Sin embargo, como $\bar{Y} = \alpha + \beta\bar{X}$,

$$E[(\alpha + \beta X - \bar{Y})^2] = E[\beta^2(X - \bar{X})^2] = \beta^2 E[(X - \bar{X})^2] \tag{2}$$

$$= \frac{\sigma_{XY}^2}{\sigma_X^4}\sigma_X^2 = \frac{\sigma_{XY}^2}{\sigma_X^2}$$

Entonces, (1) se convierte en

$$\rho^2 = \frac{\sigma_{XY}^2}{\sigma_X^2\sigma_Y^2} \quad \text{o} \quad \rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \tag{3}$$

como deseábamos demostrar. (El signo correcto de ρ está incluido en σ_{XY}.)

8.56. Con base en la tabla 8-21. *a*) Encontrar la parábola de regresión de mínimos cuadrados que ajusta los datos. *b*) Calcular los valores de regresión (que a menudo se llaman *valores de tendencia*) para los años dados y compararlos con los valores reales. *c*) Estimar la población en 1945. *d*) Estimar la población en 1960 y compararla con el valor real, 179.3. *e*) Estimar la población en 1840 y compararla con el valor real de 17.1.

Tabla 8-21

Año	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950
Población estadounidense (millones)	23.2	31.4	39.8	50.2	62.9	76.0	92.0	105.7	122.8	131.7	151.1

Fuente: Bureau of the Census.

a) Consideremos que las variables x y y denotan el año y la población, respectivamente, durante ese año. La ecuación de la parábola de mínimos cuadrados que ajusta los datos es

$$y = a + bx + cx^2 \tag{1}$$

donde a, b y c se encuentran a partir de las ecuaciones normales

$$\begin{aligned} \textstyle\sum y &= an + b\sum x + c\sum x^2 \\ \textstyle\sum xy &= a\sum x + b\sum x^2 - c\sum x^3 \\ \textstyle\sum x^2y &= a\sum x^2 + b\sum x^3 + c\sum x^4 \end{aligned} \tag{2}$$

Es conveniente ubicar el origen de modo que el año medio, 1900, corresponda a $x = 0$, y elegir una unidad para que concuerden los años 1910, 1920, 1930, 1940, 1950 y 1890, 1880, 1870, 1860, 1850 con 1, 2, 3,

$$11a + 110c = 886.8$$
$$110b = 1\,429.8$$
$$+ 1958c = 9209.0$$

A partir de la segunda ecuación en (3), $b = 13.00$; con base en la primera y la tercera ecuaciones, $a = 76.64$, $c = 0.3974$. Entonces, la ecuación que buscamos es

(4) $$y = 76.64 + 13.00x + 0.3974x^2$$

donde el origen, $x = 0$, es el 1 de julio de 1900, y la unidad de x es de 10 años.

b) Los valores de tendencia, que se obtuvieron al sustituir $x = -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5$ en (4), se muestran en la tabla 8-23 junto con los valores verdaderos. Se observa que el ajuste es bueno.

Tabla 8-22

Año	x	y	x^2	x^3	x^4	xy	x^2y
1850	−5	23.2	25	−125	625	−116.0	580.0
1860	−4	31.4	16	−64	256	−125.6	502.4
1870	−3	39.8	9	−27	81	−119.4	358.2
1880	−2	50.2	4	−8	16	−100.4	200.8
1890	−1	62.9	1	−1	1	−62.9	62.9
1900	0	76.0	0	0	0	0	0
1910	1	92.0	1	1	1	92.0	92.0
1920	2	105.7	4	8	16	211.4	422.8
1930	3	122.8	9	27	81	368.4	1105.2
1940	4	131.7	16	64	256	526.8	2107.2
1950	5	151.1	25	125	625	755.5	3777.5
	$\sum x = 0$	$\sum y =$ 886.8	$\sum x^2 =$ 110	$\sum x^3 = 0$	$\sum x^4 =$ 1 958	$\sum xy =$ 1 429.8	$\sum x^2y =$ 9 209.0

Tabla 8-23

Año	$x = -5$ 1850	$x = -4$ 1860	$x = -3$ 1870	$x = -2$ 1880	$x = -1$ 1890	$x = 0$ 1900	$x = 1$ 1910	$x = 2$ 1920	$x = 3$ 1930	$x = 4$ 1940	$x = 5$ 1950
Valor de tendencia	21.6	31.0	41.2	52.2	64.0	76.6	90.0	104.2	119.2	135.0	151.6
Valor real	23.2	31.4	39.8	50.2	62.9	76.0	92.0	105.7	122.8	131.7	151.1

c) 1945 corresponde a $x = 4.5$, para lo cual $y = 76.64 + 13.00(4.5) + 0.3974(4.5)^2 = 143.2$.

d) 1960 corresponde a $x = 6$, para lo cual $y = 76.64 + 13.00(6) + 0.3974(6)^2 = 168.9$. Lo que no concuerda muy bien con el valor real de 179.3.

e) 1840 corresponde a $x = -6$, para lo cual $y = 76.64 + 13.00(-6) + 0.3974(-6)^2 = 12.9$. Este resultado no concuerda con el valor real de 17.1.

Este ejemplo ilustra el hecho de que una relación que es satisfactoria para un rango de valores no necesariamente lo es para un rango de valores extendido.

		= –	= –			
		– – – – – –	 – – – – –		– – – – – – –	
$\sum x =$ 452.64 =	$\sum x =$ 975.16 $\bar{y} =$ 97.52			$\sum x'^2 =$ 449.38	$\sum x'y' =$ −94.67	$\sum y'^2 =$ 93.69

Entonces, mediante la fórmula de momento-producto,

$$r = \frac{\sum x'y'}{\sqrt{\left(\sum x'^2\right)\left(\sum y'^2\right)}} = \frac{-94.67}{\sqrt{(449.38)(93.69)}} = -0.4614$$

b) Concluimos que existe alguna correlación negativa entre los precios de las acciones y los bonos (esto es, que los precios de las acciones tiendan a decrecer cuando aumentan los precios de los bonos, y viceversa), aunque esta relación no es marcada.

Otro método

La tabla 8-26 muestra los rangos de los precios promedio de acciones y bonos para los años 1950 a 1959 en el orden de precios crecientes. La tabla señala también las diferencias en rango entre d y $\sum d^2$.

	–	–			–						
											$\sum d^2 =$

Entonces, $$r_{\text{rango}} = 1 - \frac{6\sum d^2}{n(n^2 - 1)} = 1 - \frac{6(272)}{10(10^2 - 1)} = -0.6485$$

Esto se compara de manera favorable con el resultado del primer método.

8.58. La tabla 8-27 muestra las distribuciones de frecuencia de las calificaciones finales de 100 estudiantes en matemáticas y física. Con referencia a esta tabla, determinar *a*) el número de estudiantes que reciben calificaciones de 70 a 79 en matemáticas y de 80 a 89 en física, *b*) el porcentaje de estudiantes con calificaciones en matemáticas menores de 70, *c*) el número de estudiantes que reciben calificaciones de 70 o más en física y menores de 80 en matemáticas, *d*) el porcentaje de estudiantes que aprueban por lo menos uno de los temas si se supone que 60 es la calificación mínima aprobatoria.

Tabla 8-27
CALIFICACIONES EN MATEMÁTICAS

CALIFICACIONES EN FÍSICA	40-49	50-59	60-69	70-79	80-89	90-99	TOTALES
90-99				2	4	4	10
80-89			1	4	6	5	16
70-79			5	10	8	1	24
60-69	1	4	9	5	2		21
50-59	3	6	6	2			17
40-49	3	5	4				12
TOTALES	7	15	25	23	20	10	100

a) Proceder hacia abajo en la columna con el encabezado 70-79 (calificación en matemáticas) hasta la fila marcada 80-89 (calificación en física). La entrada 4 proporciona el número de estudiantes.

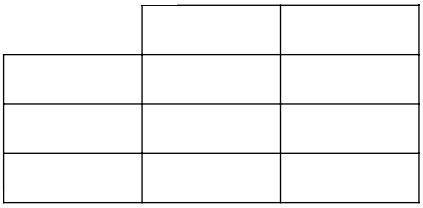
b) El número total de estudiantes con calificaciones en matemáticas menores de 70

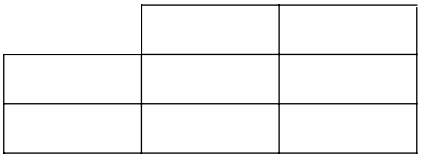
= (número con calificaciones 40-49) + (número con calificaciones 50-59)+ (número con calificaciones 60-69) = 7 + 15 + 25 = 47.

Porcentaje de estudiantes con calificaciones en matemáticas menores que 70 = 47/100 = 47%.

c) El número de estudiantes es el total que contiene la tabla 8-28, ésta es una parte de la 8-27.

Número de estudiantes = 1 + 5 + 2 + 4 + 10 =22.





$$= \frac{n\sum fxy - \left(\sum f_x x\right)\left(\sum f_y y\right)}{\sqrt{\left[n\sum f_x x^2 - \left(\sum f_x x\right)^2\right]\left[n\sum f_y y^2 - \left(\sum f_y y\right)^2\right]}}$$

Si $x = x_0 + c_x u_x$ y $y = y_0 + c_y u_y$, donde c_x y c_y son los anchos de los intervalos de clase (supuestos constantes) y x_0 y y_0 son calificaciones arbitrarias de clase que corresponden a las variables, la fórmula anterior se convierte en

$$r = \frac{n\sum f u_x u_y - \left(\sum f_x u_x\right)\left(\sum f_y u_y\right)}{\sqrt{\left[n\sum f_x u_x^2 - \left(\sum f_x u_x\right)^2\right]\left[n\sum f_y u_y^2 - \left(\sum f_y u_y\right)^2\right]}} \tag{2}$$

Éste es el *método de compilación* que se usa en el capítulo 5 como un método abreviado para calcular medias, desviaciones estándar y momentos superiores.

8.60. Encontrar el coeficiente de la correlación lineal de las calificaciones de matemáticas y física que se dan en el problema 8.58.

Con la fórmula (2) del problema 8.59 es posible ordenar el trabajo como se muestra en la tabla 8-30, que se denomina *tabla de correlación*.

Tabla 8-30

		Calificaciones en matemáticas, x									
		x	44.5	54.5	64.5	74.5	84.5	94.5	f_y	$f_y u_y$	$f_y u_y^2$	Suma de los números de la esquina en cada fila
Calificaciones en física, y	y	u_y \ u_y	−2	−1	0	1	2	3				
	94.5	2				2 [4]	4 [16]	4 [24]	10	20	40	44
	84.5	1			1 [0]	4 [4]	6 [12]	5 [15]	16	16	16	31
	74.5	0			5 [0]	10 [0]	8 [0]	1 [0]	24	0	0	0
	64.5	−1	1 [12]	4 [4]	9 [0]	5 [−5]	2 [−4]		21	−21	21	−3
	54.5	−2	3 [12]	6 [12]	6 [0]	2 [−4]			17	−34	68	20
	44.5	−3	3 [18]	5 [15]	4 [0]				12	−36	108	33
	f_y		7	15	25	23	20	10	$\Sigma f_x = \Sigma f_y = n = 100$	$\Sigma f_y u_x = -55$	$\Sigma f_y u_x^2 = 253$	$\Sigma f u_x u_y = 125$
	$f_y u_y$		−14	−15	0	23	40	30	$\Sigma f_x u_x = 64$			
	$f_y u_y^2$		28	15	0	23	80	90	$\Sigma f_x u_x^2 = 236$			
	Suma de los números de la esquina en cada columna		32	31	0	−1	24	39	$\Sigma f u_x u_y = 125$			

Verificar

El número en la esquina de cada celda representa el producto $fu_x u_y$, donde f es la frecuencia de la celda. La suma de estos números de las esquinas en cada fila se indica en la fila correspondiente de la última columna. La suma de estos números de las esquinas en cada columna se indica en la columna correspondiente de la última fila. Los totales finales de la última fila y la última columna son iguales y representan $\Sigma f u_x u_y$.

A partir de la tabla 8-30, se tiene

$$r = \frac{n\sum fu_x u_y - \left(\sum f_x u_x\right)\left(\sum f_y u_y\right)}{\sqrt{\left[n\sum f_x u_x^2 - \left(\sum f_x u_x\right)^2\right]\left[n\sum f_y u_y^2 - \left(\sum f_y u_y\right)^2\right]}}$$

$$= \frac{(100)(125) - (64)(-55)}{\sqrt{[(100)(236) - (64)^2][(100)(253) - (-55)^2]}} = \frac{16\,020}{\sqrt{(19\,504)(22\,275)}} = 0.7686$$

8.61. Usar la tabla de correlación del problema 8.60 para calcular *a*) s_x, *b*) s_y, *c*) s_{xy} y verificar la fórmula $r = s_{xy}/s_x s_y$.

a)
$$s_x = c_x\sqrt{\frac{\sum f_x u_x^2}{n} - \left(\frac{\sum f_x u_x}{n}\right)^2} = 10\sqrt{\frac{236}{100} - \left(\frac{64}{100}\right)^2} = 13.966$$

b)
$$s_y = c_y\sqrt{\frac{\sum f_y u_y^2}{n} - \left(\frac{\sum f_y u_y}{n}\right)^2} = 10\sqrt{\frac{253}{100} - \left(\frac{-55}{100}\right)^2} = 14.925$$

c)
$$s_{xy} = c_x c\left[\frac{\sum fu_x u_y}{n} - \left(\frac{\sum f_x u_x}{n}\right)\left(\frac{\sum f_y u_y}{n}\right)\right] = (10)(10)\left[\frac{125}{100} - \left(\frac{64}{100}\right)\left(\frac{-55}{100}\right)\right] = 160.20$$

Por tanto, las desviaciones estándar de las calificaciones de matemáticas y física son 14.0 y 14.9, respectivamente, mientras que su covarianza es 160.2. Tenemos

$$\frac{s_{xy}}{s_x s_y} = \frac{160.20}{(13.966)(14.925)} = 0.7686$$

concordando con el valor de r que se encontró en el problema 8.60.

8.62. Escribir las ecuaciones de las rectas de regresión de *a*) y en x y *b*) x en y para los datos del problema 8.60.

De la tabla 8-30, tenemos

$$\bar{x} = x_0 + c_x\frac{\sum f_x u_x}{n} = 64.5 + \frac{(10)(64)}{100} = 70.9$$

$$\bar{y} = y_0 + c_y\frac{\sum f_y u_y}{n} = 74.5 + \frac{(10)(-55)}{100} = 69.0$$

A partir de los resultados del problema 8.61, $s_x = 13.966$, $s_y = 14.925$ y $r = 0.7686$. Ahora se usa (16), página 268, para obtener las ecuaciones de las rectas de regresión.

a) $$y - \bar{y} = \frac{rs_y}{s_x}(x - \bar{x}), \quad y - 69.0 = \frac{(0.7686)(14.925)}{13.966}(x - 70.9),$$

o sea $$y - 69.0 = 0.821(x - 70.9)$$

b) $$x - \bar{x} = \frac{rs_x}{s_y}(y - \bar{y}), \qquad y - 70.0 = \frac{(0.7686)\,(13.966)}{14.925}(y - 69.0),$$

o sea $$x - 70.9 = 0.719(y - 69.0)$$

8.63. Calcular los errores estándar de estimación *a*) $s_{y.x}$, *b*) $s_{y.x}$ para los datos del problema 8.60. Utilizar los resultados del problema 8.61.

a) $$s_{y.x} = s_y\sqrt{1 - r^2} = 14.925\sqrt{1 - (0.7686)^2} = 9.548$$

b) $$s_{x.y} = s_x\sqrt{1 - r^2} = 13.966\sqrt{1 - (0.7686)^2} = 8.934$$

PROBLEMAS SUPLEMENTARIOS

RECTA DE MÍNIMOS CUADRADOS

8.64. Ajuste la recta de mínimos cuadrados de los datos de la tabla 8-31 usando *a*) x como la variable independiente y *b*) x como la variable dependiente. Grafique los datos y las rectas de mínimos cuadrados utilizando el mismo conjunto de ejes de coordenadas.

Tabla 8-31

x	3	5	6	8	9	11
y	2	3	4	6	5	8

8.65. Para los datos del problema 8.64, encuentre *a*) los valores de y cuando $x = 5$ y $x = 12$ y *b*) el valor de x cuando $y = 7$.

8.66. La tabla 8-32 muestra las calificaciones finales de álgebra y física que obtuvieron 10 estudiantes seleccionadas al azar de un grupo grande. *a*) Grafique los datos. *b*) Determine la recta de mínimos cuadrados que ajusta los datos, con x como la variable independiente. *c*) Calcule la recta de mínimos cuadrados que ajusta los datos, con y como

CURVAS DE REGRESIÓN DE MÍNIMOS CUADRADOS

8.68. Ajuste la parábola de mínimos cuadrados $y = a + bx + cx^2$ a los datos en la tabla 8-34.

Tabla 8-34

x	0	1	2	3	4	5	6
y	2.4	2.1	3.2	5.6	9.3	14.6	21.9

8.69. La tabla 8-35 proporciona la distancia de paro d (pies) de un automóvil que viaja a una velocidad v (millas por hora) al momento en que se detecta peligro. *a*) Grafique d contra v. *b*) Ajuste la parábola de mínimos cuadrados de la forma $d = a + bv + cv^2$ a los datos. *c*) Estime d cuando $v = 45$ millas por hora y 80 millas por hora.

Tabla 8-35

Velocidad, v (millas por hora)	20	30	40	50	60	70
Distancia de paro, d (pies)	54	90	138	206	292	396

8.70. El número y de bacterias por volumen unitario presentes en un cultivo después de x horas se indica en la tabla 8-36. *a*) Grafique los datos en papel semilogarítmico, con la escala logarítmica para y y la aritmética para x. *b*) Ajuste una curva de mínimos cuadrados de la forma $y = ab^x$ a los datos y explique por qué esta ecuación particular debe producir buenos resultados. *c*) Compare los valores de y que se obtuvieron de esta ecuación con los valores reales. *d*) Estime el valor de y cuando $x = 7$.

Tabla 8-36

Número de horas (x)	0	1	2	3	4	5	6
Número de bacterias por volumen unitario (y)	32	47	65	92	132	190	275

REGRESIÓN MÚLTIPLE

8.71. La tabla 8-37 muestra los valores que corresponden a las tres variables, x, y y z. *a*) Encuentre la ecuación de regresión lineal de mínimos cuadrados de z en x y y. *b*) Estime z cuando $x = 10$ y $y = 6$.

Tabla 8-37

ERROR ESTÁNDAR DE ESTIMACIÓN Y COEFICIENTE DE CORRELACIÓN LINEAL

8.72. Encuentre *a*) $s_{y.x}$, *b*) $s_{x.y}$ para los datos del problema 8.67.

8.73. Calcule *a*) la variación total *y*, *b*) la variación no explicada en *y*, *c*) la variación explicada en *y* para los datos del problema 8.67.

8.74. Utilice los resultados del problema 8.73 para calcular el coeficiente de correlación entre los dos conjuntos de calificaciones de pruebas del problema 8.67.

8.75. Encuentre la covarianza para los datos del problema 8.67 *a*) directamente y *b*) al emplear la fórmula $s_{xy} = rs_x s_y$ y el resultado del problema 8.74.

8.76. La tabla 8-38 muestra las edades *x* y las presiones sanguínea sistólica *y* de 12 mujeres. *a*) Calcular el coeficiente de correlación entre *x* y *y*. *b*) Determinar la recta de regresión de mínimos cuadrados de *y* en *x*. *c*) Estimar la presión sanguínea de una mujer que tiene 45 años.

Tabla 8-38

COEFICIENTE DE CORRELACIÓN GENERALIZADA

8.83. Encuentre el error estándar de estimación de z en x y y para los datos del problema 8.71.

8.84. Calcule el coeficiente de correlación múltiple para los datos del problema 8.71.

CORRELACIÓN DE RANGOS

8.85. A dos jueces les pidieron clasificar 8 candidatos que participaban en un concurso, A, B, C, D, E, F, G y H, en el orden de sus preferencias. Entregaron sus elecciones de acuerdo con la tabla 8-40. Encuentre el coeficiente de correlación de rangos y decida qué tan bien concuerdan las elecciones de los jueces.

Tabla 8-40

Candidato	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Primer juez	5	2	8	1	4	6	3	7
Segundo juez	4	5	7	3	2	8	1	6

8.86. Determine el coeficiente de correlación de rangos para los datos *a*) del problema 8.67, *b*) del problema 8.76.

8.87. Calcule el coeficiente de correlación de rangos para los datos del problema 8.82.

TEORÍA MUESTRAL DE REGRESIÓN

8.88. Con base en una muestra de tamaño 27, se encontró que una ecuación de regresión de y en x es $y = 25.0 + 2.00x$. Si $s_{y,x} = 1.50$, $s_x = 3.00$ y $\bar{x} = 7.50$, determine los límites de confianza para *a*) 95% y *b*) 99% del coeficiente de regresión.

8.89. En el problema 8.88, compruebe la hipótesis de que el coeficiente de regresión poblacional es *a*) tan bajo como 1.70, *b*) tan alto como 2.20, a un nivel de significancia de 0.01.

8.90. En el problema 8.88, calcule los límites de confianza para *a*) 95% y *b*) 99% de y cuando $x = 6.00$.

8.91. En el problema 8.88, determine los límites de confianza de *a*) 95% y *b*) 99% de la media de todos los valores de y que corresponden a $x = 6.00$.

8.92. Con referencia al problema 8.76, calcule los límites de confianza de 95% para *a*) el coeficiente de regresión de y en x, *b*) la presión sanguínea de todas las mujeres que tienen 45 años, *c*) la media de las presiones sanguíneas de todas las mujeres que tienen 45 años.

TEORÍA MUESTRAL DE CORRELACIÓN

8.93. Se calculó igual a 0.40 un coeficiente de correlación basado en una muestra de tamaño 27. ¿Es posible concluir a un nivel de significancia de *a*) 0.05 y *b*) 0.01 que el coeficiente de correlación poblacional es significativamente mayor que cero?

8.94. Un coeficiente de correlación que se basa en una muestra de tamaño 35 resultó ser 0.50. ¿Es posible rechazar la hipótesis de que el coeficiente de correlación poblacional es *a*) tan pequeño como $\rho = 0.30$ y *b*) tan grande como $\rho = 0.70$, con base en un nivel de significancia de 0.05?

8.95. Determine los límites de confianza de *a*) 95% y *b*) 99% correspondientes a un coeficiente de correlación que se calcula como 0.60 respecto de una muestra de tamaño 28.

8.96. Resuelva el problema 8.95 si el tamaño de la muestra es 52.

8.97. Calcule los límites de confianza de 95% del coeficiente de correlación que se calculó en el problema 8.76.

8.98. Dos coeficientes de correlación que se obtuvieron de muestras de tamaños 23 y 28 se calcularon, respectivamente, como 0.80 y 0.95. ¿Se puede concluir a un nivel de significancia de *a*) 0.05 y *b*) 0.01 que hay una diferencia significativa entre los dos coeficientes?

RESULTADOS DIVERSOS

8.99. Las rectas de regresión de mínimos cuadrados muestrales de un conjunto de datos que implican X y Y están dadas por $2x - 5y = 3$, $5x - 8y = 2$. Encuentre el coeficiente de correlación lineal.

8.100. Calcule el coeficiente de correlación entre las estaturas y pesos de 300 hombres adultos en Estados Unidos de acuerdo con la tabla 8-41.

Índice de precios al consumidor para atención médica = 100)	106.0	111.1	117.2	121.3	125.2	128.0	132.6	138.0

8.106. Con los datos de la tabla 8-44. *a*) Grafique los datos. *b*) Calcule la parábola de mínimos cuadrados que ajusta los datos. *c*) Calcule los valores de tendencia y compárelos con los valores reales. *d*) Explique por qué la ecuación que se obtuvo en *b*) no es útil para fines de extrapolación.

Tabla 8-44

Año	1915	1920	1925	1930	1935	1940	1945	1950	1955
Tasa de nacimiento poblacional, por 1 000	25.0	23.7	21.3	18.9	16.9	17.9	19.5	23.6	24.6

Fuente: Department of Health and Human Services.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

$y = -\frac{1}{3} + \frac{5}{7}x$ $y = -0.333 + 0.714x$ $x = 1 + \frac{9}{7}y$ o sea $x = 1.00 + 1.29y$

$= 29.13 + 0.661x$ $x = -14.39 + 1.15y$

$y = 4.000 + 0.500x$ $x = 2.408 + 0.612y$

$y = 5.51 + 3.20(x - 3) + 0.733(x - 3)^2$ o $y = 2.51 - 1.20x + 0.733x^2$

$d = 41.77 - 1.096v + 0.08786v^2$

$y = 32.14(1.427)^x$ o $y = 32.14(10)^{0.1544x}$ o $y = 32.14e^{0.3556x}$

$= 61.40 - 3.65x + 2.54y$ *b*) 40 **8.72.** *a*) 1.304 *b*) 1.443

8.73. *a*) 24.50 *b*) 17.00 *c*) 7.50 **8.74.** 0.5533 **8.75.** 1.5

8.76. *a*) 0.8961 *b*) $y = 80.78 + 1.138x$ *c*) 132

8.77. *a*) 0.958 *b*) 0.872 **8.78.** *a*) $y = 0.8x + 12$ *b*) $x = 0.45y + 1$

8.79. *a*) 1.60 *b*) 1.20 **8.80.** ± 0.80 **8.81.** 75% **8.82.** *a*) -0.9203

8.83. 3.12 **8.84.** 0.9927 **8.85.** $r_{\text{rango}} = \frac{2}{3}$ **8.86.** *a*) 0.5182 *b*) 0.9318

8.87. -1.0000 **8.88.** *a*) 2.00 ± 0.21 *b*) 2.00 ± 0.28

8.89. *a*) Usando la prueba de una cola, podemos rechazar la hipótesis.
b) Usando la prueba de una cola, no podemos rechazar la hipótesis.

8.90. *a*) 37.0 ± 3.6 *b*) 37.0 ± 4.9 **8.91.** *a*) 37.0 ± 1.5 *b*) 37.0 ± 2.1

8.92. *a*) 1.138 ± 0.398 *b*) 132.0 ± 19.2 *c*) 132.0 ± 5.4

8.93. *a*) Sí. *b*) No. **8.94.** *a*) No. *b*) Sí.

8.95. *a*) 0.2923 y 0.7951 *b*) 0.1763 y 0.8361

8.96. *a*) 0.3912 y 0.7500 *b*) 0.3146 y 0.7861

8.97. 0.7096 y 0.9653 **8.98.** *a*) sí *b*) no **8.99.** 0.8

8.100. 0.5440 **8.101.** *a*) $y = 4.44x - 142.22$ *b*) 141.9 y 177.5 libras

8.102. *a*) 16.92 1ibras *b*) 2.07 pulg **8.103.** 0.4961 y 0.7235 **8.104.** 0.9263

8.105. *b*) $y = 122.42 + 2.19x$ si x unidades es $\frac{1}{2}$ año y el origen está en 1 enero, 1954; o $y = 107.1 + 4.38x$ si x unidades es 1 año y el origen está en 1 julio, 1950.
d) 142.1 *e*) 1971

8.106. *b*) $y = 18.16 - 0.1083x + 0.4653x^2$, donde y es la tasa de nacimiento por 1 000 de población y x unidades es 5 años con origen en 1 julio, 1935.

CAPÍTULO 9

EL PROPÓSITO DEL ANÁLISIS DE VARIANZA

En el capítulo 7 se usó la teoría muestral para probar el significado de una diferencia entre dos medias muestrales. Se suponía que las dos poblaciones de las cuales se tomaron las muestras tenían la misma varianza. En muchas situaciones se necesita comprobar el significado de diferencias entre tres o más medias muestrales, o comprobar de manera equivalente la hipótesis nula de que las medias muestrales son todas iguales.

EJEMPLO 9.1 Suponga que en un experimento agrícola cuatro diferentes tratamientos químicos de tierras rinden producciones medias de trigo de 28, 22, 18 y 24 quintales por acre, respectivamente. ¿Existe una diferencia significativa entre estas medias o la gama observada se debe simplemente a la casualidad?

Problemas como el anterior pueden resolverse empleando una importante técnica que se conoce como *análisis de varianza*, desarrollada por Fisher, que usa la distribución F que se explicó en capítulos anteriores.

CLASIFICACIÓN UNIDIRECCIONAL O EXPERIMENTOS DE UN FACTOR

En un *experimento de un factor* se obtienen mediciones u observaciones de a grupos independientes de muestras, donde el número de medidas en cada grupo es b. Se habla de a *tratamientos*, cada uno con b *repeticiones* o *replicaciones*. En el ejemplo 9.1, $a = 4$.

Es posible representar los resultados de un experimento de un factor en una tabla que consta de a renglones y b columnas (tabla 9-1). Aquí, x_{jk} denota la medida del renglón j-ésimo y la columna k-ésima, donde $j = 1, 2, \ldots, a$ y $k = 1, 2, \ldots, b$. Por ejemplo, x_{35} se refiere a la quinta medida para el tercer tratamiento.

Tabla 9-1

Tratamiento 1	x_{11} x_{12} $\cdots$ x_{1b}	$\bar{x}_{1.}$
Tratamiento 2	x_{21} x_{22} $\cdots$ x_{2b}	$\bar{x}_{2.}$
$\vdots$	$\vdots$	
Tratamiento a	x_{a1} x_{a2} $\cdots$ x_{ab}	$\bar{x}_{a.}$

La media de las medidas del renglón j-ésimo se denotará por $\bar{x}_{j.}$

$$\bar{x}_{j.} = \frac{1}{b}\sum_{k=1}^{b} x_{jk} \qquad j = 1, 2, \ldots, a \tag{1}$$

El punto en $\bar{x}_{j.}$ se usa para mostrar que el índice k se ha sumado. Los valores $\bar{x}_{j.}$ se llaman *medias de grupo*, *medias de tratamiento* o *medias de renglón*. La *gran media* o *media global* es la media de todas las medidas de todos los grupos y se denota por $\bar{x}$, esto es,

$$\bar{x} = \frac{1}{ab}\sum_{j,k} x_{jk} = \frac{1}{ab}\sum_{j=1}^{a}\sum_{k=1}^{b} x_{jk} \tag{2}$$

VARIACIÓN TOTAL. VARIACIÓN DENTRO DE LOS TRATAMIENTOS. VARIACIÓN ENTRE TRATAMIENTOS

La *variación total*, que se denota por v, se define como la suma de los cuadrados de las desviaciones de cada medida de la gran media $\bar{x}$, esto es,

$$\text{Variación total} = v = \sum_{j,k}(x_{jk} - \bar{x})^2 \tag{3}$$

Al escribir la identidad

$$x_{jk} - \bar{x} = (x_{jk} - \bar{x}_{j.}) + (\bar{x}_{j.} - \bar{x}) \tag{4}$$

y entonces elevarla al cuadrado y sumarla sobre j y k, se demuestra (vea el problema 9.1) que

$$\sum_{j,k}(x_{jk} - \bar{x})^2 = \sum_{j,k}(x_{jk} - \bar{x}_{j.})^2 + \sum_{j,k}(\bar{x}_{j.} - \bar{x})^2 \tag{5}$$

o bien

$$\sum_{j,k}(x_{jk} - \bar{x})^2 = \sum_{j,k}(x_{jk} - \bar{x}_{j.})^2 + b\sum_{j}(\bar{x}_{j.} - \bar{x})^2 \tag{6}$$

La primera sumatoria a la derecha de (5) o (6) se llama *variación con tratamientos* (puesto que implica los cuadrados de las desviaciones de x_{jk} de las medias de tratamiento $\bar{x}_{j.}$) y se denota por v_w. Por tanto,

$$v_w = \sum_{j,k}(x_{jk} - \bar{x}_{j.})^2 \tag{7}$$

La segunda sumatoria a la derecha de (5) o (6) se llama *variación entre tratamientos* (puesto que implica los cuadrados de las desviaciones de las medias de los diversos tratamientos $\bar{x}_{j.}$ de la gran media $\bar{x}$) y se denota por v_b. En consecuencia,

$$v_b = \sum_{j,k}(\bar{x}_{j.} - \bar{x})^2 = b\sum_{j}(\bar{x}_{j.} - \bar{x})^2 \tag{8}$$

Las ecuaciones (5) y (6) se pueden reescribir como

$$v = v_w + v_b \tag{9}$$

MÉTODOS CORTOS PARA OBTENER VARIACIONES

Para minimizar el trabajo de calcular las variaciones anteriores, resultan convenientes las siguientes fórmulas:

$$v = \sum_{j,k}x_{jk}^2 - \frac{\tau^2}{ab} \tag{10}$$

$$v_b = \frac{1}{b}\sum_{j}\tau_{j.}^2 - \frac{\tau^2}{ab} \tag{11}$$

$$v_w = v - v_b \tag{12}$$

donde τ es el total de todos los valores x_{jk} y $\tau_{j.}$ es el total de todos los valores del tratamiento j-ésimo. Es decir,

$$\tau = \sum_{j,k}x_{jk} \qquad \tau_{j.} = \sum_{k}x_{jk} \tag{13}$$

En la práctica es conveniente restar algún valor fijo de todos los datos de la tabla; lo anterior no tiene ningún efecto en los resultados finales.

MODELO MATEMÁTICO LINEAL PARA ANÁLISIS DE VARIANZA

Es posible considerar que cada renglón de la tabla 9-1 representa una muestra aleatoria de tamaño b de la población para ese tratamiento particular. Por tanto, para el tratamiento j se tienen las variables aleatorias independientes y dis-

tribuidas de manera idéntica, $X_{j1}, X_{j2}, \ldots, X_{jb}$, las cuales toman, respectivamente, los valores de $x_{j1}, x_{j2}, \ldots, x_{jb}$. Cada una de las X_{jk} ($k = 1, 2, \ldots, b$) puede expresarse como la suma de su valor esperado y un término de "aleatoriedad" o "error":

$$X_{jk} = \mu_j + \Delta_{jk} \tag{14}$$

Las Δ_{jk} pueden tomarse como variables aleatorias independientes (relativas tanto a j como a k) y normalmente distribuidas con media cero y varianza σ^2. Esto es equivalente a suponer que las X_{jk} ($j = 1, 2, \ldots, a$; $k = 1, 2, \ldots, b$) son variables normales mutuamente independientes con medias μ_j y una varianza común σ^2.

Se define la variable μ por

$$\mu = \frac{1}{a}\sum_j \mu_j$$

Se puede pensar en μ como la media para un tipo de población grande que abarca el tratamiento de todas las poblaciones. Entonces, (14) puede reescribirse como (vea el problema 9.18)

$$X_{jk} = \mu + \alpha_j + \Delta_{jk} \quad \text{donde} \quad \sum_j \alpha_j = 0 \tag{15}$$

La constante α_j puede verse como el efecto especial del tratamiento j-ésimo.

La hipótesis nula de que todas las medias de los tratamiento son iguales está dada por (H_0: $\alpha_j = 0; j = 1, 2, \ldots, a$) o de manera equivalente por (H_0: $\mu_j = \mu; j = 1, 2, \ldots, a$). Si H_0 es verdadera, las poblaciones del tratamiento, que se suponen normales, tienen una media común además de una varianza común. En consecuencia, existe sólo una población en tratamiento y todos los tratamientos son estadísticamente idénticos.

VALORES ESPERADOS DE LAS VARIACIONES

Las variaciones V_b entre tratamientos, la variación V_w interna de los tratamientos y la variación total V son variables aleatorias que, respectivamente, tienen los valores v_b, v_w, y v se definen en (8), (7) y (3). Se puede demostrar (problema 9.19) que

$$E(V_b) = (a - 1)\sigma^2 + b\sum_j \alpha_j^2 \tag{16}$$

$$E(V_w) = a(b - 1)\sigma^2 \tag{17}$$

$$E(V) = (ab - 1)\sigma^2 + b\sum \alpha_j^2 \tag{18}$$

De (17) se deduce que

$$E\left[\frac{V_w}{a(b - 1)}\right] = \sigma^2 \tag{19}$$

de modo que

$$\hat{S}_w^2 = \frac{V_w}{a(b - 1)} \tag{20}$$

siempre es una mejor estimación (insesgada) de σ^2 sin importar si H_0 es verdadera o no. Por otro lado, de (16) y (18) se observa que sólo si H_0 es verdadera se tendrá

$$E\left(\frac{V_b}{a - 1}\right) = \sigma^2 \qquad E\left(\frac{V}{ab - 1}\right) = \sigma^2 \tag{21}$$

por lo que sólo en tal caso

$$\hat{S}_b^2 = \frac{V_b}{a - 1} \qquad \hat{S}^2 = \frac{V}{ab - 1} \tag{22}$$

proporcionarán estimaciones insesgadas de σ^2. Pero si H_0 no es verdad, entonces de (16) se tiene

$$E(\hat{S}_b^2) = \sigma^2 + \frac{b}{a-1}\sum_j \alpha_j^2 \tag{23}$$

DISTRIBUCIONES DE LAS VARIACIONES

Empleando el teorema 4-4, página 115, se pueden probar los siguientes teoremas fundamentales acerca de las distribuciones de las variaciones V_w, V_b y V.

Teorema 9-1 V_w/σ^2 tiene una distribución ji cuadrada con $a(b-1)$ grados de libertad.

Teorema 9-2 Con base en la hipótesis nula, H_0, V_b/σ^2 y V/σ^2 tienen distribuciones ji cuadrada con $a-1$ y $ab-1$ grados de libertad, respectivamente.

Es importante enfatizar que el teorema 9-1 es válido si se supone o no H_0, mientras que el 9-2 es válido sólo si se supone H_0.

LA PRUEBA *F* PARA LA HIPÓTESIS NULA DE MEDIAS IGUALES

Si la hipótesis nula H_0 no es verdadera, esto es, si las medias del tratamiento no son iguales, de (23) se puede esperar que $\hat{S}_b^2$ sea mayor que σ^2, y el efecto se vuelve más pronunciado a medida que aumenta la discrepancia entre las medias. Por otro lado, a partir de (19) y (20) se espera que $\hat{S}_w^2$ sea igual a σ^2 sin importar si las medias son iguales o no. Se concluye que un buen estadístico para probar la hipótesis H_0 lo proporciona $\hat{S}_b^2/\hat{S}_w^2$. Si éste es significativamente grande, es posible concluir que existe una diferencia significativa entre las medias de tratamiento, por lo que se rechaza H_0. De lo contrario, se puede aceptar H_0 o dejar el juicio pendiente para un análisis posterior.

Para usar este estadístico, es necesario conocer su distribución, la cual se proporciona en el siguiente teorema, que es una consecuencia del teorema 5-8, página 159.

Teorema 9-3 El estadístico $F = \hat{S}_b^2/\hat{S}_w^2$ tiene la distribución F con $a-1$ y $a(b-1)$ grados de libertad.

El teorema 9-3 permite probar la hipótesis nula a un nivel de significancia especificado mediante el empleo de una prueba de una cola de la distribución F.

TABLAS DE ANÁLISIS DE VARIANZA

Los cálculos que se requieren para la prueba anterior se resumen en la tabla 9-2, que se llama *tabla de análisis de varianza*. En la práctica, se deberían calcular v y v_b utilizando el método largo (3) y (8), o el método corto (10) y (11), y después se debería calcular $v_w = v - v_b$. Se debe notar que los grados de libertad para la variación total, esto es, $ab-1$, es igual a la suma de los grados de libertad para las variaciones entre tratamientos y dentro de tratamientos.

Tabla 9-2

Variación	Grados de libertad	Cuadrado medio	F
Entre tratamientos, $v_b = b\sum_j(\bar{x}_{j.} - \bar{x})^2$	$a-1$	$\hat{s}_b^2 = \dfrac{v_b}{a-1}$	$\dfrac{\hat{s}_b^2}{\hat{s}_w^2}$ con $a-1, a(b-1)$ grados de libertad
Dentro de tratamientos, $v_w = v - v_b$	$a(b-1)$	$\hat{s}_w^2 = \dfrac{v_w}{a(b-1)}$	
Total, $v = v_b + v_w = \sum_{j,k}(x_{jk} - \bar{x})^2$	$ab-1$		

MODIFICACIONES PARA NÚMEROS DESIGUALES DE OBSERVACIONES

En el caso de que los tratamientos 1, . . . , a tengan diferentes números de observaciones iguales, respectivamente, a $n_1, \ldots, n_a$, los resultados anteriores se modifican con facilidad. En consecuencia se obtiene,

$$v = \sum_{j,k}(x_{jk} - \bar{x})^2 = \sum_{j,k}x_{jk}^2 - \frac{\tau^2}{n} \tag{24}$$

$$v_b = \sum_{j,k}(\bar{x}_{j.} - \bar{x})^2 = \sum_{j}n_j(\bar{x}_{j.} - \bar{x})^2 = \sum_{j}\frac{\tau_{j.}^2}{n_j} - \frac{\tau^2}{n} \tag{25}$$

$$v_w = v - v_b \tag{26}$$

donde $\sum_{j,k}$ denota la sumatoria sobre k de 1 a n_j y después sobre j de 1 hasta a, $n = \sum_j n_j$ es el número total de observaciones en todos los tratamientos, τ es la suma de todas las observaciones, $\tau_{j.}$ es la suma de todos los valores en el tratamiento j-ésimo, y $\sum_j$ es la suma de $j = 1$ hasta a.

La tabla de análisis de varianza para este caso está dada en la tabla 9-3.

Tabla 9-3

Variación	Grados de libertad	Cuadrado medio	F
Entre tratamientos, $v_b = \sum_j n_j(\bar{x}_{j.} - \bar{x})^2$	$a - 1$	$\hat{s}_b^2 = \frac{v_b}{a-1}$	$\frac{\hat{s}_b^2}{\hat{s}_w^2}$ con $a - 1, n - a$ grados de libertad
Dentro de tratamientos, $v_w = v - v_b$	$n - a$	$\hat{s}_w^2 = \frac{v_w}{n-a}$	
Total, $v = v_b + v_w = \sum_{j,k}(x_{jk} - \bar{x})^2$	$n - 1$		

CLASIFICACIÓN BIDIRECCIONAL O EXPERIMENTOS DE DOS FACTORES

Pueden generalizarse las ideas del análisis de varianza para la clasificación unidireccional o los experimentos de un factor. Se ilustrará el procedimiento para la *clasificación bidireccional* o *experimento de dos factores*.

EJEMPLO 9.2 Suponga que un experimento agrícola consiste en examinar la producción por acre de 4 variedades de trigo, donde cada variedad se cultiva en 5 lotes de terreno distintos. Entonces, se necesita un total de (4)(5) = 20 lotes. En tal caso, es conveniente combinar lotes en *bloques*, digamos, 4 lotes por bloque, con variedades de trigo que se cultivan en cada lote dentro de un bloque. Por este motivo, aquí serían necesarios 5 bloques.

En este caso existen dos clasificaciones o factores, puesto que es posible que haya diferencias en la producción por acre debidas a *i*) el tipo particular de trigo que se cultiva, o *ii*) el bloque particular que se usa (lo cual puede implicar diferente fertilidad en la tierra, etcétera).

Por analogía con el experimento agrícola del ejemplo 9.2, a menudo se hace referencia a las dos clasificaciones o factores de un experimento como *tratamientos* y *bloques*, pero claramente es posible denominar simplemente como el factor 1 y el factor 2, etcétera.

NOTACIÓN PARA EXPERIMENTOS DE DOS FACTORES

Suponiendo que se tienen a tratamientos y b bloques, se construye la tabla 9-4, donde se supone que existe un valor experimental (por ejemplo, producción por acre) correspondiente a cada tratamiento o bloque. Para el tratamiento j y el bloque k, este valor se denota por x_{jk}. La media de las entradas en el renglón j-ésimo se denota mediante $\bar{x}_{j.}$, donde $j = 1, \ldots, a$,

mientras que la media de las entradas en la columna k-ésima se denota mediante $\bar{x}_{.k}$, donde $k = 1, \ldots, b$. La *gran media* o *media global* se denota por medio de $\bar{x}$. En símbolos:

$$\bar{x}_{j.} = \frac{1}{b}\sum_{k=1}^{b} x_{jk}, \qquad \bar{x}_{.k} = \frac{1}{a}\sum_{j=1}^{a} x_{jk}, \qquad \bar{x} = \frac{1}{ab}\sum_{j,k} x_{jk} \tag{27}$$

Tabla 9-4

Bloques

Tratamientos	1	2	...	b	
1	x_{11}	x_{12}	$\cdots$	x_{1b}	$\bar{x}_{1.}$
2	x_{21}	x_{22}	$\cdots$	x_{2b}	$\bar{x}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots\vdots\vdots$		
a	x_{a1}	x_{a2}	$\cdots$	x_{ab}	$\bar{x}_{a.}$
	$\bar{x}_{.1}$	$\bar{x}_{.2}$	$\cdots$	$\bar{x}_{.b}$	

VARIACIONES PARA EXPERIMENTOS DE DOS FACTORES

Como en el caso de experimentos de un factor, se pueden definir las variaciones para experimentos de dos factores. Primero, se define la *variación total*, como en (3), del siguiente modo

$$v = \sum_{j,k}(x_{jk} - \bar{x})^2 \tag{28}$$

Al escribir la identidad

$$x_{jk} - \bar{x} = (x_{jk} - \bar{x}_{j.} - \bar{x}_{.k} + \bar{x}) + (\bar{x}_{j.} - \bar{x}) + (\bar{x}_{.k} - \bar{x}) \tag{29}$$

y entonces, elevando al cuadrado y sumando después sobre j y k, se demuestra que

$$v = v_e + v_r + v_c \tag{30}$$

donde v_e = variación debida a error o al azar $= \sum_{j,k}(x_{jk} - \bar{x}_{j.} - \bar{x}_{.k} + \bar{x})^2$

$$v_r = \text{variación entre renglones (tratamientos)} = b\sum_{j=1}^{a}(\bar{x}_{j.} - \bar{x})^2$$

$$v_c = \text{variación entre columnas (bloques)} = a\sum_{k=1}^{b}(\bar{x}_{.k} - \bar{x})^2$$

La variación debida al error o al azar también se conoce como *variación residual*.

A continuación se muestran fórmulas cortas para el cálculo, que son análogas a (10), (11) y (12).

$$v = \sum_{j,k} x_{jk}^2 - \frac{\tau^2}{ab} \tag{31}$$

$$v_r = \frac{1}{b}\sum_{j=1}^{a}\tau_{j.}^2 - \frac{\tau^2}{ab} \tag{32}$$

$$v_c = \frac{1}{a}\sum_{k=1}^{b}\tau_{.k}^2 - \frac{\tau^2}{ab} \tag{33}$$

$$v_e = v - v_r - v_c \tag{34}$$

donde $\tau_{j.}$ es el total de entradas del renglón j-ésimo, $\tau_{.k}$ es el total de entradas en la columna k-ésima, y τ es el total de todas las entradas.

ANÁLISIS DE LA VARIANZA PARA EXPERIMENTOS DE DOS FACTORES

En el caso del modelo matemático de experimentos de dos factores, se supone que las variables aleatorias X_{jk}, cuyos valores son x_{jk}, pueden escribirse como

$$X_{jk} = \mu + \alpha_j + \beta_k + \Delta_{jk} \tag{35}$$

Aquí μ es la gran media poblacional, α_j es la parte de X_{jk} que se debe a los tratamientos diferentes (que a veces se denominan *efectos de tratamiento*), β_k es la parte de X_{jk} debida a bloques diferentes (llamados algunas veces *efectos de bloques*) y Δ_{jk} es la parte de X_{jk} que se debe al azar o error. Como ya se observó, es posible tomar las Δ_{jk} como las variables aleatorias, independientes y distribuidas de manera normal con media cero y varianza σ^2, de modo que X_{jk} son también variables libres y distribuidas de manera normal con varianza σ^2. Con suposiciones adecuadas sobre las medias de X_{jk}, se tiene

$$\sum_j \alpha_j = 0 \qquad \sum_k \beta_k = 0 \tag{36}$$

lo que hace

$$\mu = \frac{1}{ab}\sum_{j,k} E(X_{jk})$$

Correspondiendo a los resultados (16) a (18), es posible demostrar que

$$E(V_r) = (a - 1)\sigma^2 + b\sum_j \alpha_j^2 \tag{37}$$

$$E(V_c) = (b - 1)\sigma^2 + a\sum_k \beta_k^2 \tag{38}$$

$$E(V_e) = (a - 1)(b - 1)\sigma^2 \tag{39}$$

$$E(V) = (ab - 1)\sigma^2 + b\sum_j \alpha_j^2 + a\sum_k \beta_k^2 \tag{40}$$

Existen dos hipótesis nulas que se desearían probar:

$H_0^{(1)}$: Todos los tratamientos (renglones) tienen medias iguales, es decir, $\alpha_j = 0, j = 1, \ldots, a$

$H_0^{(2)}$: Todos los bloques (columnas) tienen medias iguales, es decir, $\beta_k = 0, k = 1, \ldots, b$

Se observa en (39) que, sin importar $H_0^{(1)}$ o $H_0^{(2)}$, una mejor estimación (insesgada) de σ^2 está dada por

$$\hat{S}_e^2 = \frac{V_e}{(a - 1)(b - 1)} \qquad \text{es decir,} \qquad E(\hat{S}_e^2) = \sigma^2 \tag{41}$$

Además, si las hipótesis $H_0^{(1)}$ y $H_0^{(2)}$ son verdaderas, entonces

$$\hat{S}_r^2 = \frac{V_r}{a - 1}, \quad \hat{S}_c^2 = \frac{V_c}{b - 1}, \qquad \hat{S}^2 = \frac{V}{ab - 1} \tag{42}$$

serán estimaciones insesgadas de σ^2. No obstante, si $H_0^{(1)}$ y $H_0^{(2)}$ no son verdaderas, a partir de (37) y (38), respectivamente, se tiene

$$E(\hat{S}_r^2) = \sigma^2 + \frac{b}{a - 1}\sum_j \alpha_j^2 \tag{43}$$

$$E(\hat{S}_c^2) = \sigma^2 + \frac{a}{b - 1}\sum_k \beta_k^2 \tag{44}$$

Los siguientes teoremas son similares a los teoremas 9-1 y 9-2.

Teorema 9-4 V_e/σ^2 es una distribución ji cuadrada con $(a - 1)(b - 1)$ grados de libertad, sin importar $H_0^{(1)}$ o $H_0^{(2)}$.

Teorema 9-5 Con base en la hipótesis $H_0^{(1)}$, V_r/σ^2 es una distribución ji cuadrada con $a-1$ grados de libertad. Con base en la hipótesis $H_0^{(2)}$, V_c/σ^2 es una distribución ji cuadrada con $b-1$ grados de libertad. Con base en las dos hipótesis $H_0^{(1)}$ y $H_0^{(2)}$, V/σ^2 es una distribución ji cuadrada con $ab-1$ grados de libertad.

Para probar la hipótesis $H_0^{(1)}$ es natural considerar el estadístico $\hat{S}_r^2/\hat{S}_e^2$, puesto que se observa de (43) que se espera que $\hat{S}_r^2$ difiera de manera significativa de σ^2 si las medias del renglón (tratamiento) son significativamente diferentes. De manera similar, para probar la hipótesis $H_0^{(2)}$, se considera el estadístico $\hat{S}_c^2/\hat{S}_e^2$. Las distribuciones de $\hat{S}_r^2/\hat{S}_e^2$ y $\hat{S}_c^2/\hat{S}_e^2$ están dadas en el siguiente teorema, análogo al 9-3.

Teorema 9-6 Con base en la hipótesis $H_0^{(1)}$ el estadístico $\hat{S}_r^2/\hat{S}_e^2$ tiene una distribución F con $a-1$ y $(a-1)(b-1)$ grados de libertad. Con base en la hipótesis $H_0^{(2)}$, el estadístico $\hat{S}_c^2/\hat{S}_e^2$ tiene distribución F con $b-1$ y $(a-1)(b-1)$ grados de libertad.

El teorema permite aceptar o rechazar $H_0^{(1)}$ y $H_0^{(2)}$ a niveles de significancia específicos. Por conveniencia, como en el caso de un factor, una tabla de análisis de varianza se construye como se muestra en la tabla 9-5.

Tabla 9-5

Variación	Grados de libertad	Cuadrado medio	F
Entre tratamientos, $v_r = b\sum_j(\bar{x}_{j.} - \bar{x})^2$	$a-1$	$\hat{s}_r^2 = \dfrac{v_r}{a-1}$	$\hat{s}_r^2/\hat{s}_e^2$ con $a-1$ $(a-1)(b-1)$ grados de libertad
Entre bloques, $v_c = a\sum_k(\bar{x}_{.k} - \bar{x})^2$	$b-1$	$\hat{s}_c^2 = \dfrac{v_r}{b-1}$	$\hat{s}_c^2/\hat{s}_e^2$ con $b-1$ $(a-1)(b-1)$ grados de libertad
Residual o aleatoria, $v_e = v - v_r - v_c$	$(a-1)(b-1)$	$\hat{s}_e^2 = \dfrac{v_e}{(a-1)(b-1)}$	
Total, $v = v_r + v_c + v_e = \sum_{j,k}(x_{jk} - \bar{x})^2$	$ab-1$		

EXPERIMENTOS DE DOS FACTORES CON REPLICACIÓN

En la tabla 9-4 existe sólo una entrada correspondiente a un tratamiento y a un bloque particular. A menudo se puede obtener más información si se repite el experimento, proceso que se denomina *replicación*, en cuyo caso habrá más de una entrada que corresponda a un tratamiento y a un bloque particular. Se supondrá que existen c entradas para cada posición; es posible hacer cambios adecuados cuando los números replicados no son todos iguales.

Debido a la replicación, es necesario usar un modelo adecuado para sustituir al que está dado por (35), página 320. Para obtener este resultado, se deja que X_{jkl} denote la variable aleatoria correspondiente al j-ésimo renglón o tratamiento, la columna o bloque k-ésimo y la repetición o replicación, l-ésima. El modelo está dado por

$$X_{jkl} = \mu + \alpha_j + \beta_k + \gamma_{jk} + \Delta_{jkl}$$

donde μ, α_j, β_k, se definen como antes, las Δ_{jkl} son variables aleatorias independientes y distribuidas de manera normal con media cero y varianza σ^2, mientras que γ_{jk} representan los *efectos de interacción* de renglón-columna o tratamiento-bloque (que, a menudo, se llaman *interacciones*). Con respecto a (36) se tiene

$$\sum_j \alpha_j = 0, \qquad \sum_k \beta_k = 0, \qquad \sum_j \gamma_{jk} = 0, \qquad \sum_k \gamma_{jk} = 0 \tag{45}$$

Como antes, la variación total v de todos los datos puede descomponerse en variaciones debidas a renglones v_r, columnas v_c y error aleatorio o residual v_e:

$$v = v_r + v_c + v_i + v_e \tag{46}$$

donde

$$v = \sum_{j,k,l}(x_{jkl} - \bar{x})^2 \tag{47}$$

$$v_r = bc\sum_{j=1}^{a}(\bar{x}_{j..} - \bar{x})^2 \tag{48}$$

$$v_c = ac\sum_{k=1}^{b}(\bar{x}_{.k.} - \bar{x})^2 \tag{49}$$

$$v_i = c\sum_{j,k}(\bar{x}_{jk.} - \bar{x}_{j..} - \bar{x}_{.k.} + \bar{x})^2 \tag{50}$$

$$v_e = \sum_{j,k,l}(x_{jkl} - \bar{x}_{jk.})^2 \tag{51}$$

En estos resultados, los puntos en los subíndices tienen significados análogos a los que se dieron antes (página 319). Por ejemplo,

$$\bar{x}_{j..} = \frac{1}{bc}\sum_{k,l}x_{jkl} = \frac{1}{b}\sum_{k}\bar{x}_{jk.} \tag{52}$$

Utilizando un número adecuado de grados de libertad (gl) para cada fuente de variación, es posible establecer la tabla del análisis de variación, tabla 9-6.

Tabla 9-6

Variación	Grados de libertad	Cuadrado medio	F
Entre tratamientos, v_r	$a - 1$	$\hat{s}_r^2 = \dfrac{v_r}{a - 1}$	$\hat{s}_r^2/\hat{s}_e^2$ con $a - 1$, $ab(c - 1)$ grados de libertad
Entre bloques, v_c	$b - 1$	$\hat{s}_e^2 = \dfrac{v_c}{b - 1}$	$\hat{s}_c^2/\hat{s}_e^2$ con $b - 1$, $ab(c - 1)$ grados de libertad
Interacción, v_i	$(a - 1)(b - 1)$	$\hat{s}_i^2 = \dfrac{v_i}{(a - 1)(b - 1)}$	$\hat{s}_i^2/\hat{s}_e^2$ con $(a - 1)(b - 1)$, $ab(c - 1)$ grados de libertad
Residual o aleatoria, v_e	$ab(c - 1)$	$\hat{s}_e^2 = \dfrac{v_e}{ab(c - 1)}$	
Total, v	$abc - 1$		

Pueden usarse los cocientes F en la última columna de la tabla 9-6 para probar la hipótesis nula

$H_0^{(1)}$: Todos los tratamientos (renglones) tienen medias iguales, es decir, $\alpha_j = 0$

$H_0^{(2)}$: Todos los bloques (columnas) tienen medias iguales, es decir, $\beta_k = 0$

$H_0^{(3)}$: No hay interacción entre tratamientos y bloques, es decir, $\gamma_{jk} = 0$

Desde una perspectiva práctica, primero se debe decidir si se puede o no rechazar $H_0^{(3)}$ al nivel de significancia adecuado utilizando la razón $F\,\hat{s}_i^2/\hat{s}_e^2$ de la tabla 9-6. En ese caso, surgen dos posibilidades.

Caso I *No se puede rechazar* $H_0^{(3)}$: En este caso, se concluye que las interacciones no son demasiado grandes. Entonces, se puede probar $H_0^{(1)}$ y $H_0^{(2)}$ mediante los cocientes F: $\hat{s}_r^2/\hat{s}_e^2$ y $\hat{s}_c^2/\hat{s}_e^2$, respectivamente, como se muestra en la tabla 9-6. Algunos estadísticos recomiendan juntar las variaciones de este caso, tomando el total $v_i + v_e$ y dividiéndolo entre los grados de libertad totales correspondientes $(a - b)(b - 1) + ab(c - 1)$, y luego usar este valor para sustituir el denominador $\hat{s}_e^2$ en la prueba F.

Caso II *Se puede rechazar* $H_0^{(3)}$: En este caso se concluye que las interacciones son significativamente grandes. Las diferencias en factores serían entonces importantes sólo si fueran grandes en comparación con tales interacciones. Por esta razón, muchos estadísticos recomiendan que $H_0^{(1)}$ y $H_0^{(2)}$ se prueben usando las razones F: $\hat{s}_r^2/\hat{s}_i^2$ y $\hat{s}_c^2/\hat{s}_i^2$ en vez de los que están en la tabla 9-6. Se usará también este procedimiento alternativo.

El análisis de varianza con replicación se emplea con mayor facilidad si primero se totalizan los valores de replicación que corresponden a tratamientos (renglones) y a bloques (columnas) particulares. Esta operación produce una tabla de dos factores con entradas únicas, que se pueden analizar como en la tabla 9-5. El procedimiento se ilustra en el problema 9.13.

DISEÑO EXPERIMENTAL

Las técnicas de análisis de varianza que ya se explicaron anteriormente, se emplean después de que se obtienen los resultados de un experimento. No obstante, para lograr tanta información como sea posible, es necesario planear con antelación y con cuidado los detalles de un experimento, lo que a menudo se conoce como *diseño del experimento*. A continuación se dan ejemplos importantes de diseño experimental.

1. **ALEATORIZACIÓN COMPLETA.** Suponga que se pretende realizar un experimento agrícola como en el ejemplo 9.1, página 314. Para diseñar un experimento de este tipo, se divide el terreno en $4 \times 4 = 16$ lotes (como se indica en la figura 9-1 con cuadros, aunque físicamente cualquier forma sirve) y asignar a cada tratamiento, indicado por A, B, C y D, a cuatro bloques que se eligen de manera completamente aleatoria. El propósito de la aleatorización es eliminar diversas fuentes de error, como la fertilidad de la tierra.

D	A	C	C
B	D	B	A
D	C	B	D
A	B	C	A

Aleatorización completa
Figura 9-1

I	C	B	A	D
II	A	B	D	C
III	B	C	D	A
IV	A	D	C	B

Aleatorización de bloques
Figura 9-2

D	B	C	A
B	D	A	C
C	A	D	B
A	C	B	D

Cuadrado latino
Figura 9-3

B_γ	A_β	D_δ	C_α
A_δ	B_α	C_γ	D_β
D_α	C_δ	B_β	A_γ
C_β	D_γ	A_α	B_δ

Cuadrado grecolatino
Figura 9-4

2. **ALEATORIZACIÓN DE BLOQUES.** Como en el ejemplo 9.2, cuando es necesario tener un conjunto completo de tratamientos para cada bloque, los tratamientos A, B, C y D se introducen en orden aleatorio dentro de cada bloque I, II, III, IV (vea la figura 9-2) y, por tanto, los bloques se llaman *bloques aleatorizados*. Este tipo de diseño se usa cuando se desea controlar *una fuente de error o variabilidad*, o sea, la diferencia entre los bloques (los renglones de la figura 9-2).

3. **CUADRADOS LATINOS.** Para algunos propósitos, es necesario controlar *dos fuentes de error o variabilidad* al mismo tiempo, como la diferencia de los renglones y la de las columnas. En el experimento del ejemplo 9.1, por ejemplo, es posible que los errores en diferentes renglones y columnas se deban a cambios en la fertilidad de la tierra en diferentes partes del terreno. En ese caso, es deseable que cada tratamiento ocurra una vez en cada renglón y una vez en cada columna, como en la figura 9-3. El arreglo se llama *cuadrado latino* porque se usan las letras latinas A, B, C, D.

4. CUADRADOS GRECOLATINOS. Si es necesario controlar *tres* fuentes de error o variabilidad, se usa un *cuadrado grecolatino*, como se indica en la figura 9-4. Este cuadrado consiste básicamente en dos cuadrados latinos superpuestos, usando las letras latinas A, B, C, D para un cuadrado mientras que se utilizan las letras griegas α, β, γ, δ para los otros cuadrados. El requisito adicional que debe cumplirse es que se necesita usar cada letra latina sólo una vez con cada letra griega. Cuando se cumple esta propiedad, se dice que el cuadrado es *ortogonal.*

PROBLEMAS RESUELTOS

CLASIFICACIÓN UNIDIRECCIONAL O EXPERIMENTOS DE UN FACTOR

9.1. Demostrar que

$$\sum_{j,k}(x_{jk} - \bar{x})^2 = \sum_{j,k}(x_{jk} - \bar{x}_{j.})^2 + \sum_{j,k}(\bar{x}_{j.} - \bar{x})^2$$

Se tiene $x_{jk} - \bar{x} = (x_{jk} - \bar{x}_{j.}) + (\bar{x}_{j.} - \bar{x})$. Entonces, elevando al cuadrado y sumando sobre j y k, encontramos

$$\sum_{j,k}(x_{jk} - \bar{x})^2 = \sum_{j,k}(x_{jk} - \bar{x}_{j.})^2 + \sum_{j,k}(\bar{x}_{j.} - \bar{x})^2 + 2\sum_{j,k}(x_{jk} - \bar{x}_{j.})(\bar{x}_{j.} - \bar{x})$$

Para comprobar el resultado es necesario demostrar que la última sumatoria es cero. Para hacer esto, procedemos como se muestra a continuación.

$$\begin{aligned}\sum_{j,k}(x_{jk} - \bar{x}_{j.})(\bar{x}_{j.} - \bar{x}) &= \sum_{j=1}^{a}(\bar{x}_{j.} - \bar{x})\left[\sum_{k=1}^{b}(x_{jk} - \bar{x}_{j.})\right] \\ &= \sum_{j=1}^{a}(\bar{x}_{j.} - \bar{x})\left[\left(\sum_{k=1}^{b}x_{jk}\right) - b\bar{x}_{j.}\right] = 0\end{aligned}$$

puesto que $\bar{x}_{j.} = \frac{1}{b}\sum_{k=1}^{b}x_{jk}$.

9.2. Verificar que *a*) $\tau = ab\bar{x}$, *b*) $\tau_{j.} = b\bar{x}_{j.}$, *c*) $\sum_j \tau_{j.} = ab\bar{x}$, usando la notación de la página 315.

a)
$$\tau = \sum_{j,k}x_{jk} = ab\left(\frac{1}{ab}\sum_{j,k}x_{jk}\right) = ab\bar{x}$$

b)
$$\tau_{j.} = \sum_{k}x_{jk} = b\left(\frac{1}{b}\sum_{k}x_{jk}\right) = b\bar{x}_{j.}$$

c) Puesto que $\tau_{j.} = \sum_k x_{jk}$, se tiene

$$\sum_j \tau_{j.} = \sum_j\sum_k x_{jk} = \tau = ab\bar{x}$$

donde se utilizó el resultado del inciso *a*).

9.3. Verificar las fórmulas cortas (10) a (12), página 315.

Tenemos

$$\begin{aligned}v &= \sum_{j,k}(x_{jk} - \bar{x})^2 = \sum_{j,k}\left(x_{jk}^2 - 2\bar{x}\bar{x}_{jk} + \bar{x}^2\right) \\ &= \sum_{j,k}x_{jk}^2 - 2\bar{x}\sum_{j,k}x_{jk} + ab\bar{x}^2 \\ &= \sum_{j,k}x_{jk}^2 - 2\bar{x}(ab\bar{x}) + ab\bar{x}^2 \\ &= \sum_{j,k}x_{jk}^2 - ab\bar{x}^2 \\ &= \sum_{j,k}x_{jk}^2 - \frac{\tau^2}{ab}\end{aligned}$$

utilizando el problema 9.2*a*) en las líneas tercera y última anteriores. De modo similar

$$\begin{aligned} v_b &= \sum_{j,k}(\bar{x}_{j.} - \bar{x})^2 = \sum_{j,k}\left(\bar{x}_{j.}^2 - 2\bar{x}\bar{x}_{j.} + \bar{x}^2\right) \\ &= \sum_{j,k}\bar{x}_{j.}^2 - 2\bar{x}\sum_{j,k}\bar{x}_{j.} + ab\bar{x}^2 \\ &= \sum_{j,k}\left(\frac{\tau_{j.}}{b}\right)^2 - 2\bar{x}\sum_{j,k}\frac{\tau_{j.}}{b} + ab\bar{x}^2 \\ &= \frac{1}{b^2}\sum_{j=1}^{a}\sum_{k=1}^{b}\tau_{j.}^2 - 2\bar{x}(ab\bar{x}) + ab\bar{x}^2 \\ &= \frac{1}{b}\sum_{j=1}^{a}\tau_{j.}^2 - ab\bar{x}^2 \\ &= \frac{1}{b}\sum_{j=1}^{a}\tau_{j.}^2 - \frac{\tau^2}{ab} \end{aligned}$$

usando el problema 9.2*b*) en la tercera línea y el problema 9.2*a*) en la última línea.

Finalmente, se llega a (12) a partir del hecho de que $v = v_b + v_w$ o $v_w = v - v_b$.

9.4. La tabla 9-7 muestra las producciones en quintales por acre de una variedad particular de trigo que se cultiva en un tipo específico de tierra tratada con químicos *A*, *B* o *C*. Calcular *a*) la media de las producciones para los diferentes tratamientos, *b*) la gran media para todos los tratamientos, *c*) la variación total, *d*) la variación entre tratamientos y *e*) la variación dentro de tratamientos. Utilizar el método largo.

Tabla 9-7

A	48	49	50	49
B	47	49	48	48
C	49	51	50	50

Tabla 9-8

3	4	5	4
2	4	3	3
4	6	5	5

Para simplificar la aritmética, se puede restar un número adecuado, digamos, 45, de todos los datos sin que se afecten los valores de las variaciones. Se obtienen entonces los datos de la tabla 9-8.

a) Las medias de tratamiento (renglón) para la tabla 9-8 están dadas, respectivamente, por

$$\bar{x}_{1.} = \frac{1}{4}(3 + 4 + 5 + 4) = 4, \qquad \bar{x}_{2.} = \frac{1}{4}(2 + 4 + 3 + 3) = 3, \qquad \bar{x}_{3.} = \frac{1}{4}(4 + 6 + 5 + 5) = 5$$

Por tanto, las producciones medias, obtenidas al sumar 45 a ellas, son 49, 48 y 50 quintales por acre para *A*, *B* y *C*, respectivamente.

b) $$\bar{x} = \frac{1}{12}(3 + 4 + 5 + 4 + 2 + 4 + 3 + 3 + 4 + 6 + 5 + 5) = 4$$

En consecuencia, la gran media para el conjunto de datos original es $45 + 4 = 49$ quintales por acre.

c) $$\begin{aligned} \text{Variación total} = v &= \sum_{j,k}(x_{jk} - \bar{x})^2 \\ &= (3 - 4)^2 + (4 - 4)^2 + (5 - 4)^2 + (4 - 4)^2 \\ &\quad + (2 - 4)^2 + (4 - 4)^2 + (3 - 4)^2 + (3 - 4)^2 \\ &\quad + (4 - 4)^2 + (6 - 4)^2 + (5 - 4)^2 + (5 - 4)^2 \\ &= 14 \end{aligned}$$

d) $$\begin{aligned} \text{Variación entre tratamientos} = v_b &= b\sum_{j}(\bar{x}_{j.} - \bar{x})^2 \\ &= 4[(4 - 4)^2 + (3 - 4)^2 + (5 - 4)^2] = 8 \end{aligned}$$

e) $$\text{Variación dentro de tratamientos} = v_w = v - v_b = 14 - 8 = 6$$

Otro método

$$\begin{aligned} v_w &= \sum_{j,k}(x_{jk} - \bar{x}_{j\cdot})^2 \\ &= (3-4)^2 + (4-4)^2 + (5-4)^2 + (4-4)^2 \\ &\quad + (2-3)^2 + (4-3)^2 + (3-3)^2 + (3-3)^2 \\ &\quad + (4-5)^2 + (6-5)^2 + (5-5)^2 + (5-5)^2 \\ &= 6 \end{aligned}$$

9.5. Con referencia al problema 9.4, realizar una estimación insesgada de la varianza poblacional σ^2 a partir de *a*) la variación entre tratamientos con base en la hipótesis nula de medias de tratamientos iguales, *b*) la variación dentro de tratamientos.

a) $$\hat{s}_b^2 = \frac{v_b}{a-1} = \frac{8}{3-1} = 4$$

b) $$\hat{s}_w^2 = \frac{v_w}{a(b-1)} = \frac{6}{3(4-1)} = \frac{2}{3}$$

9.6. Con referencia al problema 9.4, ¿es posible rechazar la hipótesis nula de medias iguales a un nivel de significancia de *a*) 0.05, *b*) 0.01?

Tenemos $$F = \frac{\hat{s}_b^2}{\hat{s}_w^2} = \frac{4}{2/3} = 6$$

con $a - 1 = 3 - 1 = 2$ y $a(b-1) = 3(4-1) = 9$ grados de libertad.

a) De acuerdo con el apéndice F, con $\nu_1 = 2$ y $\nu_2 = 9$, se ve que $F_{0.95} = 4.26$. Puesto que $F = 6 > F_{0.95}$, se puede rechazar la hipótesis nula de medias iguales a un nivel de significancia de 0.05.

b) De acuerdo con el apéndice F, con $\nu_1 = 2$ y $\nu_2 = 9$, se ve que $F_{0.99} = 8.02$. Puesto que $F = 6 < F_{0.99}$, no es posible rechazar la hipótesis nula de medias iguales a un nivel de significancia de 0.01.

La tabla de análisis de varianza para los problemas 9.4 a 9.6 se muestra en la tabla 9-9.

Tabla 9-9

Variación	Grados de libertad	Cuadrado medio	F
Entre tratamientos, $v_b = 8$	$a - 1 = 2$	$\hat{s}_b^2 = \frac{8}{2} = 4$	$F = \frac{\hat{s}_b^2}{\hat{s}_w^2} = \frac{4}{2/3} = 6$ con 2, 9 grados de libertad
Dentro de tratamientos, $v_w = v - v_b = 14 - 8 = 6$	$a(b-1) = (3)(3) = 9$	$\hat{s}_w^2 = \frac{6}{9} = \frac{2}{3}$	
Total, $v = 14$	$ab - 1 = (3)(4) - 1 = 11$		

9.7. Emplear las fórmulas cortas (10) a (12) para obtener los resultados del problema 9.4.

a) Tenemos

$$\sum_{j,k} x_{jk}^2 = 9 + 16 + 25 + 16 + 4 + 16 + 9 + 9 + 16 + 36 + 25 + 25 = 206$$

También

$$\tau = 3 + 4 + 5 + 4 + 2 + 4 + 3 + 3 + 4 + 6 + 5 + 5 = 48$$

Por tanto,
$$v = \sum_{j,k} x_{jk}^2 - \frac{\tau^2}{ab}$$
$$= 206 - \frac{(48)^2}{(3)(4)} = 206 - 192 = 14$$

b) Los totales de los renglones son
$$\tau_{1.} = 3 + 4 + 5 + 4 = 16$$
$$\tau_{2.} = 2 + 4 + 3 + 3 = 12$$
$$\tau_{3.} = 4 + 6 + 5 + 5 = 20$$

Además,
$$\tau = 16 + 12 + 20 = 48$$

Entonces,
$$v_b = \frac{1}{b}\sum_j \tau_{j.}^2 - \frac{\tau^2}{ab}$$
$$= \frac{1}{4}(16^2 + 12^2 + 20^2) - \frac{(48)^2}{(3)(4)} = 200 - 192 = 8$$

c)
$$v_w = v - v_b = 14 - 8 = 6$$

Es conveniente ordenar los datos como se muestra en la tabla 9-10.

Tabla 9-10

					$\tau_{j.}$	$\tau_{j.}^2$
A	3	4	5	4	16	256
B	2	4	3	3	12	144
C	4	6	5	5	20	400
	$\sum_{j,k} x_{jk}^2 = 206$				$\tau = \sum_j \tau_{j.} = 48$	$\sum_j \tau_{j.}^2 = 800$

$$v = 206 - \frac{(48)^2}{(3)(4)} = 14$$
$$v_b = \frac{1}{4}(800) - \frac{(48)^2}{(3)(4)} = 8$$

Los resultados concuerdan con los que se obtuvieron en el problema 9.4 y a partir de este punto el análisis procede como en el caso anterior.

9.8. Una compañía desea comprar una de cinco diferentes máquinas *A*, *B*, *C*, *D*, *E*. En un experimento diseñado para decidir si existe una diferencia en el desempeño entre las máquinas, cada uno de cinco operadores experimentados trabajarán en ellas durante tiempos iguales. La tabla 9-11 muestra el número de unidades que produjeron. Probar la hipótesis de que no existe ninguna diferencia entre las máquinas a un nivel de significancia de *a*) 0.05, *b*) 0.01.

Tabla 9-11

A	68	72	75	42	53
B	72	52	63	55	48
C	60	82	65	77	75
D	48	61	57	64	50
E	64	65	70	68	53

Tabla 9-12

						$\tau_{j\cdot}$	$\tau_{j\cdot}^2$
A	8	12	15	−18	−7	10	100
B	12	–8	3	−5	−2	0	0
C	0	22	6	17	15	60	3 600
D	−12	1	−3	4	−10	–20	400
E	4	5	10	8	−7	20	400
	$\sum x_{jk}^2 = 2\,356$					70	4 500

Restar un número adecuado, digamos, 60, de todos los datos para obtener la tabla 9-12.

Entonces,

$$v = 2\,356 - \frac{(70)^2}{(5)(4)} = 2\,356 - 245 = 2\,111$$

$$v_b = \frac{1}{5}(4\,500) - \frac{(70)^2}{(5)(4)} = 900 - 245 = 655$$

Ahora se forma la tabla 9-13.

Tabla 9-13

Variación	Grados de libertad	Cuadrado medio	*F*
Entre tratamientos, $v_c = 655$	$a - 1 = 4$	$\hat{s}_b^2 = \frac{655}{4} = 163.75$	$F = \frac{\hat{s}_b^2}{\hat{s}_w^2} = 2.25$
Dentro de tratamientos, $v_w = 1\,456$	$a(b-1) = 5(4) = 20$	$\hat{s}_w^2 = \frac{1\,456}{(5)(4)} = 72.8$	
Total, $v = 2\,111$	$ab - 1 = 24$		

Para 4, con 20 grados de libertad se tiene $F_{0.95} = 2.87$. Por tanto, no se puede rechazar la hipótesis nula a un nivel de significancia de 0.05, de modo que no es posible rechazarla a un nivel de 0.01.

MODIFICACIONES PARA NÚMEROS DESIGUALES DE OBSERVACIONES

9.9. La tabla 9-14 contiene los tiempos de vida en horas de muestras tomadas de tres tipos de tubos de televisores diferentes fabricados por una compañía. Empleando el método largo, probar a un nivel de significancia de *a*) 0.05, *b*) 0.01 si existe una diferencia entre los tres tipos.

Tabla 9-14

Muestra 1	407	411	409		
Muestra 2	404	406	408	405	402
Muestra 3	410	408	406	408	

Tabla 9-15

						Total	Media
Muestra 1	7	11	9			27	9
Muestra 2	4	6	8	5	2	25	5
Muestra 3	10	8	6	8		32	8
	$\bar{x}$ = gran media $= \frac{84}{12} = 7$						

Es conveniente restar un número adecuado, digamos, 400, con lo que se obtiene la tabla 9-15, en donde se indican los totales de los renglones, las medias de la muestra o grupo, y la gran media. Entonces, tenemos

$$v = \sum_{j,k}(x_{jk} - \bar{x})^2 = (7 - 7)^2 + (11 - 7)^2 + \cdots + (8 - 7)^2 = 72$$

$$v_b = \sum_{j,k}(\bar{x}_{j.} - \bar{x})^2 = \sum_{j} n_j(\bar{x}_{j.} - \bar{x})^2$$

$$= 3(9 - 7)^2 + 5(7 - 5)^2 + 4(8 - 7)^2 = 36$$

$$v_w = v - v_b = 72 - 36 = 36$$

Se puede obtener también v_w directamente al observar que es equivalente a

$$(7 - 9)^2 + (11 - 9)^2 + (9 - 9)^2 + (4 - 5)^2 + (6 - 5)^2 + (8 - 5)^2 + (5 - 5)^2$$
$$+ (2 - 5)^2 + (10 - 8)^2 + (8 - 8)^2 + (6 - 8)^2 + (8 - 8)^2$$

Los datos se resumen en la tabla de análisis de varianza 9-16.

Tabla 9-16

Variación	Grados de libertad	Cuadrado medio	F
$v_b = 36$	$a - 1 = 2$	$\hat{s}_b^2 = \frac{36}{2} = 18$	$\frac{\hat{s}_b^2}{\hat{s}_w^2} = \frac{18}{4}$ $= 4.5$
$v_w = 36$	$n - a = 9$	$\hat{s}_w^2 = \frac{36}{9} = 4$	

Ahora, para 2 y 9 grados de libertad, se encuentra del apéndice F que $F_{0.95} = 4.26$, $F_{0.99} = 8.02$. Por tanto, es posible rechazar la hipótesis de medias iguales (esto es, no existe diferencia significativa entre los tres tipos de tubos) a un nivel de 0.05 pero no al de 0.01.

9.10. Resolver el problema 9.9 empleando las fórmulas cortas incluidas en (24), (25) y (26).

A partir de la tabla 9-15,

$$n_1 = 3, \quad n_2 = 5, \quad n_3 = 4, \quad n = 12, \quad \tau_{1.} = 27, \quad \tau_{2.} = 25, \quad \tau_{3.} = 32, \quad \tau = 84$$

Por tanto, se tiene,

$$v = \sum_{j,k} x_{jk}^2 - \frac{\tau^2}{n} = 7^2 + 11^2 + \cdots + 6^2 + 8^2 - \frac{(84)^2}{12} = 72$$

$$v_b = \sum_{j} \frac{\tau_{j.}^2}{n_j} - \frac{\tau^2}{n} = \frac{(27)^2}{3} + \frac{(25)^2}{5} + \frac{(32)^2}{4} - \frac{(84)^2}{12} = 36$$

$$v_w = v - v_b = 36$$

Utilizando lo anterior, aplicamos el análisis de varianza como en el problema 9.9.

CLASIFICACIÓN BIDIRECCIONAL O EXPERIMENTOS DE DOS FACTORES

9.11. La tabla 9-17 muestra las producciones por acre de cuatro cosechas diferentes cultivadas en lotes tratados con tres tipos de fertilizante. Utilizando el método largo probar a un nivel de significancia de 0.01 si *a*) existe una diferencia significativa en la producción por acre debida a los fertilizantes, *b*) si hay una diferencia en la producción por acre debida a las cosechas.

Tabla 9-17

	Cosecha I	Cosecha II	Cosecha III	Cosecha IV
Fertilizante A	4.5	6.4	7.2	6.7
Fertilizante B	8.8	7.8	9.6	7.0
Fertilizante C	5.9	6.8	5.7	5.2

Calcular tanto los totales y las medias de los renglones como los totales y las medias de las columnas y la gran media, como se muestra en la tabla 9-18.

Tabla 9-18

	Cosecha I	Cosecha II	Cosecha III	Cosecha IV	Totales de renglón	Medias de renglón
Fertilizante A	4.5	6.4	7.2	6.7	24.8	6.2
Fertilizante B	8.8	7.8	9.6	7.0	33.2	8.3
Fertilizante C	5.9	6.8	5.7	5.2	23.6	5.9
Totales de columnas	19.2	21.0	22.5	18.9	Gran total = 81.6	
Medias de columna	6.4	7.0	7.5	6.3	Gran media = 6.8	

v_r = variación de medias de renglón de la gran media

$= 4[(6.2 - 6.8)^2 + (8.3 - 6.8)^2 + (5.9 - 6.8)^2] = 13.68$

v_c = variación de las medias de las columnas de la gran media

$= 3[(6.4 - 6.8)^2 + (7.0 - 6.8)^2\,(7.5 - 6.8)^2 + (6.3 - 6.8)^2] = 2.82$

v = variación total

$= (4.5 - 6.8)^2 + (6.4 - 6.8)^2 + (7.2 - 6.8)^2 + (6.7 - 6.8)^2$

$+ (8.8 - 6.8)^2 + (7.8 - 6.8)^2 + (9.6 - 6.8)^2 + (7.0 - 6.8)^2$

$+ (5.9 - 6.8)^2 + (6.8 - 6.8)^2 + (5.7 - 6.8)^2 + (5.2 - 6.8)^2$

$= 23.08$

v_e = variación aleatoria $= v - v_r - v_c = 6.58$

Ello conduce al análisis de varianza de la tabla 9-19.

A un nivel de significancia de 0.05 con 2, y 6 grados de libertad, $F_{0.95} = 5.14$. Entonces, puesto que $6.24 > 5.14$, se puede rechazar la hipótesis de que las medias de los renglones son iguales y concluir a un nivel de significancia de 0.05 que la diferencia en producción se debe a los fertilizantes.

Puesto que el valor *F* correspondiente a las diferencias en las medias de la columna es menor que 1, es posible concluir que no existe una diferencia significativa en producción debida a las cosechas.

Tabla 9-19

Variación	Grados de libertad	Cuadrado medio	F
$v_r = 13.68$	2	$\hat{s}_r^2 = 6.84$	$F = \hat{s}_r^2/\hat{s}_e^2 = 6.24$ gl: 2, 6
$v_c = 2.82$	3	$\hat{s}_c^2 = 0.94$	$F = \hat{s}_c^2/\hat{s}_e^2 = 0.86$ gl: 3, 6
$v_e = 6.58$	6	$\hat{s}_e^2 = 1.097$	
$v = 23.08$	11		

9.12. Emplear las fórmulas de cálculo cortas para obtener los resultados del problema 9.11.

De la tabla 9-18 se tiene

$$\sum_{j,k} x_{jk}^2 = (4.5)^2 + (6.4)^2 + \cdots + (5.2)^2 = 577.96$$

$$\tau = 24.8 + 33.2 + 23.6 = 8.16$$

$$\sum \tau_{j.}^2 = (24.8)^2 + (33.2)^2 + (23.6)^2 = 2\,274.24$$

$$\sum \tau_{.k}^2 = (19.2)^2 + (21.0)^2 + (22.5)^2 + (18.9)^2 = 1\,673.10$$

Entonces,

$$v = \sum_{j,k} x_{jk}^2 - \frac{\tau^2}{ab} = 577.96 - 554.88 = 23.08$$

$$v_r = \frac{1}{b}\sum \tau_{j.}^2 - \frac{\tau^2}{ab} = \frac{1}{4}(2\,274.24) - 554.88 = 13.68$$

$$v_c = \frac{1}{a}\sum \tau_{.k}^2 - \frac{\tau^2}{ab} = \frac{1}{3}(1\,673.10) - 554.88 = 2.82$$

$$v_e = v - v_r - v_c = 23.08 - 13.68 - 2.82 = 6.58$$

en concordancia con el problema 9.11.

EXPERIMENTOS DE DOS FACTORES CON REPLICACIÓN

9.13. Un fabricante quiere determinar la eficacia de cuatro tipos de máquinas, A, B, C, D en la producción de pernos. Para lograrlo, se obtiene el número de pernos defectuosos producidos por cada máquina en los días de cierta semana en cada uno de dos turnos. Los resultados se indican en la tabla 9-20. Llevar a cabo un análisis de varianza para probar a un nivel de significancia de 0.05 si existe *a*) una diferencia entre las máquinas, *b*) una diferencia entre turnos.

Tabla 9-20

	PRIMER TURNO					SEGUNDO TURNO				
	Lun	Mar	Miér	Jue	Vie	Lun	Mar	Miér	Jue	Vie
A	6	4	5	5	4	5	7	4	6	8
B	10	8	7	7	9	7	9	12	8	8
C	7	5	6	5	9	9	7	5	4	6
D	8	4	6	5	5	5	7	9	7	10

Los datos se organizan de manera equivalente como se muestra en la tabla 9-21. Esta tabla indica los dos factores principales, es decir, *Máquina* y *Turno*. Observe que para cada máquina están indicados dos turnos. Los días de la semana pueden considerarse como replicaciones o repeticiones del desempeño de cada máquina en los dos turnos.

Tabla 9-21

FACTOR I	FACTOR II	REPLICACIONES					
Máquina	Turno	Lun	Mar	Miér	Jue	Vie	TOTALES
A	1	6	4	5	5	4	24
	2	5	7	4	6	8	30
B	1	10	8	7	7	9	41
	2	7	9	12	8	8	44
C	1	7	5	6	5	9	32
	2	9	7	5	4	6	31
D	1	8	4	6	5	5	28
	2	5	7	9	7	10	38
	TOTALES	57	51	54	47	59	268

La variación total para todos los datos de la tabla 9-21 es

$$v = 6^2 + 4^2 + 5^2 + \cdots + 7^2 + 10^2 - \frac{(268)^2}{40} = 1\,946 - 1\,795.6 = 150.4$$

Para considerar los dos factores principales, *Máquina* y *Turno*, la atención se limita al total de los valores de replicación para cada combinación de factores. Éstos están ordenados en la tabla 9-22, por lo que es una tabla de dos factores con entradas únicas.

Tabla 9-22

	Primer turno	Segundo turno	TOTALES
A	24	30	54
B	41	44	85
C	32	31	63
D	28	38	66
TOTALES	125	143	268

La variación total de la tabla 9-22, que se llamará *variación del subtotal* v_s, está dada por

$$v_s = \frac{(24)^2}{5} + \frac{(41)^2}{5} + \frac{(32)^2}{5} + \frac{(28)^2}{5} + \frac{(30)^2}{5} + \frac{(44)^2}{5} + \frac{(31)^2}{5} + \frac{(38)^2}{5} - \frac{(268)^2}{40}$$
$$= 1\,861.2 - 1\,795.6 = 65.6$$

La variación entre los renglones es

$$v_r = \frac{(54)^2}{10} + \frac{(85)^2}{10} + \frac{(63)^2}{10} + \frac{(66)^2}{10} - \frac{(268)^2}{40} = 1\,846.6 - 1\,795.6 = 51.0$$

La variación entre las columnas corresponde a

$$v_c = \frac{(125)^2}{20} + \frac{(143)^2}{20} - \frac{(268)^2}{40} = 1\,803.7 - 1\,795.6 = 8.1$$

Si ahora se resta de la variación del subtotal v_s la suma de las variaciones entre los renglones y las columnas ($v_r + v_c$), se obtiene la variación debida a la *interacción* entre renglones y columnas, que está dada por

$$v_i = v_s - v_r - v_c = 65.6 - 51.0 - 8.1 = 6.5$$

Finalmente, la variación residual, que se puede considerar como la variación aleatoria o de error v_e (siempre que se suponga que los diversos días de la semana no proporcionan ninguna diferencia importante), se encuentra al restar la suma de las variaciones de renglón, columna, e interacción (esto es, la variación del subtotal) de la variación total v, lo que produce

$$v_e = v - (v_r + v_c + v_i) = v - v_s = 150.4 - 65.6 = 84.8$$

Estas variaciones se indican en el análisis de varianza, en la tabla 9-23. Ésta también proporciona el número de grados de libertad correspondientes a cada tipo de variación. Por tanto, puesto que hay 4 renglones en la tabla 9-22, la variación que se debe a los renglones tiene $4 - 1 = 3$ grados de libertad, mientras que la variación debida a las 2 columnas tiene $2 - 1 = 1$ grados de libertad. Para determinar los grados de libertad debidos a la interacción, se debe observar que hay 8 entradas en la tabla 9-22. Por tanto, el total de los grados de libertad es $8 - 1 = 7$. Puesto que 3 de éstos se deben a los renglones y 1 a las columnas, el restante, $7 - (3 + 1) = 3$, se debe a la interacción. Puesto que hay 40 entradas en la tabla original 9-21, el total de grados de libertad es $40 - 1 = 39$. En consecuencia, los grados de libertad que se deben a la variación aleatoria o residual son $39 - 7 = 32$.

Tabla 9-23

Variación	Grados de libertad	Cuadrado medio	F
Renglones (máquinas), $v_r = 51.0$	3	$\hat{s}_r^2 = 17.0$	$\frac{17.0}{2.65} = 6.42$
Columna (turnos) $v_c = 8.1$	1	$\hat{s}_c^2 = 8.1$	$\frac{8.1}{2.65} = 3.06$
Interacción, $v_i = 6.5$	3	$\hat{s}_i^2 = 2.167$	$\frac{2.167}{2.65} = 0.817$
Subtotal, $v_s = 65.6$	7		
Aleatorio o residual, $v_e = 84.8$	32	$\hat{s}_e^2 = 2.65$	
Total, $v = 150.4$	39		

Para seguir adelante, primero es necesario determinar si existe una interacción significativa entre los factores básicos (esto es, los renglones y las columnas de la tabla 9-22). A partir de la tabla 9-23, se observa que en la interacción $F = 0.817$, que indica que la interacción no es significativa, esto es, no se puede rechazar la hipótesis $H_0^{(3)}$ de la página 323. De acuerdo con las reglas en la página 323, se nota que la F calculada en los renglones es 6.42. Puesto que $F_{0.95} = 2.90$ para 3, 32 grados de libertad, es posible rechazar la hipótesis $H_0^{(1)}$ de que los renglones tienen medias iguales. Esto equivale a decir que a un nivel de 0.05, se concluye que las máquinas no tienen la misma eficacia.

Para 1, 32 grados de libertad $F_{0.95} = 4.15$. Entonces, puesto que la F calculada para las columnas es 3.06, *no se puede* rechazar la hipótesis $H_0^{(2)}$ de que las columnas tienen medias iguales. Esto equivale a decir que a un nivel de 0.05 no hay una diferencia significativa entre los turnos.

Si se elige analizar los resultados y se agrupan las interacciones y las variaciones residuales como recomiendan algunos estadísticos, se encuentra para la variación agrupada y los grados de libertad (gl) agrupados $v_i + v_e = 6.5 + 84.8 = 91.3$ y $3 + 32 = 35$, respectivamente, lo que conduce a una varianza agrupada de $91.3 / 35 = 2.61$. El uso de este valor en vez de 2.65 en el denominador de F en la tabla 9-23 no afecta las conclusiones anteriores.

9.14. Resolver el problema 9.13 a un nivel de 0.01.

A este nivel, todavía no existe una interacción apreciable, por esto procederemos así.

Puesto que $F_{0.99} = 4.47$ para 3, 32 grados de libertad, y como la F que se calculó en los renglones es 6.42, se puede concluir que, incluso a un nivel de 0.01, las máquinas no tienen la misma eficacia.

Puesto que $F_{0.99} = 7.51$ para 1, 32 grados de libertad, y puesto que la F calculada para las columnas es 3.06, se puede concluir al nivel de 0.01 que no existen diferencias significativas entre los turnos.

CUADRADOS LATINOS

9.15. Un campesino desea probar los efectos de cuatro diferentes fertilizantes, A, B, C, D sobre la producción de trigo. Para eliminar las fuentes de error debidas a la variabilidad en la fertilidad, usa los fertilizantes en un cuadrado latino arreglado como se muestra en la tabla 9-24, donde los números indican las producciones en quintales por área unitaria. Llevar a cabo un análisis de varianza para determinar si hay una diferencia significativa entre los fertilizantes a los niveles de significancia de *a*) 0.05, *b*) 0.01.

Tabla 9-24

A 18	*C* 21	*D* 25	*B* 11
D 22	*B* 12	*A* 15	*C* 19
B 15	*A* 20	*C* 23	*D* 24
C 22	*D* 21	*B* 10	*A* 17

Tabla 9-25

					TOTALES
	A 18	*C* 21	*D* 25	*B* 11	75
	D 22	*B* 12	*A* 15	*C* 19	68
	B 15	*A* 20	*C* 23	*D* 24	82
	C 22	*D* 21	*B* 10	*A* 17	70
TOTALES	77	74	73	71	295

Tabla 9-26

	A	*B*	*C*	*D*	
TOTAL	70	48	85	92	295

Primero se obtienen los totales para los renglones y las columnas como se indica en la tabla 9-25. Luego se obtienen las producciones totales de cada uno de los fertilizantes, como se muestra en la tabla 9-26. Después, la variación total y las variaciones para los renglones, las columnas y los tratamientos se obtienen como de costumbre. Se encuentra que

$$\text{Variación total} = v = (18)^2 + (21)^2 + (25)^2 + \cdots + (10)^2 + (17)^2 - \frac{(295)^2}{16}$$
$$= 5\,769 - 5\,439.06 = 329.94$$

$$\text{Variación entre renglones} = v_r = \frac{(75)^2}{4} + \frac{(68)^2}{4} + \frac{(82)^2}{4} + \frac{(70)^2}{4} - \frac{(295)^2}{16}$$
$$= 5\,468.25 - 5\,439.06 = 29.19$$

$$\text{Variación entre columnas} = v_c = \frac{(77)^2}{4} + \frac{(74)^2}{4} + \frac{(73)^2}{4} + \frac{(71)^2}{4} - \frac{(295)^2}{16}$$
$$= 5\,443.75 - 5\,439.06 = 4.69$$

$$\text{Variación entre tratamientos} = v_t = \frac{(70)^2}{4} + \frac{(48)^2}{4} + \frac{(85)^2}{4} + \frac{(92)^2}{4} - \frac{(295)^2}{16}$$
$$= 5\,723.25 - 5\,439.06 = 284.19$$

El análisis de varianza ahora se muestra en la tabla 9-27.

a) Puesto que $F_{0.95,3,6} = 4.76$, a un nivel de significancia de 0.05 es posible rechazar la hipótesis de que existen medias de renglones iguales. Se deduce que a un nivel de significancia de 0.05 existe una diferencia en la fertilidad de la tierra entre los renglones.

Puesto que el valor de F para la columna es menor que 1, se concluye que no hay diferencia entre la fertilidad de la tierra en las columnas.

Puesto que el valor de F para los tratamientos es $47.9 > 4.76$, se concluye que sí hay una diferencia entre los fertilizantes.

Tabla 9-27

Variación	Grados de libertad	Cuadrado medio	F
Renglones, 29.19	3	9.73	4.92
Columnas, 4.69	3	1.563	0.79
Tratamientos, 284.19	3	94.73	47.9
Residuales, 11.87	6	1.978	
Total, 329.94	15		

b) Puesto que $F_{0.99,3,6} = 9.78$, se puede aceptar la hipótesis de que no hay diferencia en la fertilidad de la tierra entre los renglones (o la columnas) a un nivel de significancia de 0.01. Sin embargo, es necesario sostener la conclusión de que hay una diferencia entre fertilizantes a un nivel de 0.01.

CUADRADOS GRECOLATINOS

9.16. Es de interés determinar si existe una diferencia entre el número de millas por galón que se logra con las gasolinas *A*, *B*, *C*, *D.* Diseñar un experimento que incluya cuatro diferentes conductores, coches y carreteras.

Puesto que número (cuatro) de gasolinas, conductores, coches y carreteras es el mismo, es posible usar un *cuadrado grecolatino*. Se supone que los coches se representan por los renglones y los conductores por las columnas, como se indica en la tabla 9-28. Ahora se asignan las diferentes gasolinas *A*, *B*, *C*, *D* a renglones y columnas de manera aleatoria, lo que se sujeta sólo al requisito de que cada letra aparece sólo una vez en cada renglón y sólo una vez en cada columna. En consecuencia, cada conductor tendrá la oportunidad de manejar cada coche y de usar cada tipo de gasolina (ningún coche utiliza dos veces la misma gasolina).

Ahora se asignan, de manera aleatoria, las cuatro carreteras que se van a utilizar, denotadas por α, β, γ, δ, y sujetas al mismo requisito impuesto a las letras latinas. Por tanto, cada conductor tendrá una oportunidad de manejar en cada una de las carreteras. Un posible arreglo está dado en la tabla 9-28.

Tabla 9-28

		CONDUCTORES			
		1	**2**	**3**	**4**
COCHES	**1**	B_γ	A_β	D_δ	C_α
	2	A_α	B_α	C_γ	D_β
	3	D_α	C_δ	B_β	A_γ
	4	C_β	D_γ	A_α	B_δ

9.17. Suponer que, al llevar a cabo el experimento del problema 9.16, los números de millas por galón están dados como se muestra en la tabla 9-29. Emplear el análisis de varianza para determinar si hay alguna diferencia significativa a un nivel de 0.05.

Primero obtenemos los totales de los renglones y de las columnas como se muestra en la tabla 9-30.

Tabla 9-29

		CONDUCTORES			
		1	**2**	**3**	**4**
COCHES	**1**	B_γ 19	A_β 16	D_δ 16	C_α 14
	2	A_δ 15	B_α 18	C_γ 11	D_β 15
	3	D_α 14	C_δ 11	B_β 21	A_γ 16
	4	C_β 16	D_γ 16	A_α 15	B_δ 23

Tabla 9-30

					TOTALES
	B_γ 19	A_β 16	D_δ 16	C_α 14	65
	A_δ 15	B_α 18	C_γ 11	D_β 15	59
	D_α 14	C_δ 11	B_β 21	A_γ 16	62
	C_β 16	D_γ 16	A_α 15	B_δ 23	70
TOTALES	64	61	63	68	256

A continuación se obtienen los totales para cada letra latina y para cada letra griega:

$$A \text{ total: } 15 + 16 + 15 + 16 = 62$$
$$B \text{ total: } 19 + 18 + 21 + 23 = 81$$
$$C \text{ total: } 16 + 11 + 11 + 14 = 52$$
$$D \text{ total: } 14 + 16 + 16 + 15 = 61$$
$$\alpha \text{ total: } 14 + 18 + 15 + 14 = 61$$
$$\beta \text{ total: } 16 + 16 + 21 + 15 = 68$$
$$\gamma \text{ total: } 19 + 16 + 11 + 16 = 62$$
$$\delta \text{ total: } 15 + 11 + 16 + 23 = 65$$

Ahora calculamos las variaciones que corresponden a todo lo anterior, utilizando el método corto.

$$\text{Renglones: } \frac{(65)^2}{4} + \frac{(59)^2}{4} + \frac{(62)^2}{4} + \frac{(70)^2}{4} - \frac{(256)^2}{16} = 4\,112.50 - 4\,096 = 16.50$$

$$\text{Columnas: } \frac{(64)^2}{4} + \frac{(61)^2}{4} + \frac{(63)^2}{4} + \frac{(68)^2}{4} - \frac{(256)^2}{16} = 4\,102.50 - 4\,096 = 6.50$$

$$\text{Gasolinas } (A, B, C, D): \frac{(62)^2}{4} + \frac{(81)^2}{4} + \frac{(52)^2}{4} + \frac{(61)^2}{4} - \frac{(256)^2}{16} = 4\,207.50 - 4\,096 = 111.50$$

$$\text{Caminos } (\alpha, \beta, \gamma, \delta): \frac{(61)^2}{4} + \frac{(68)^2}{4} + \frac{(62)^2}{4} + \frac{(65)^2}{4} - \frac{(256)^2}{16} = 4\,103.50 - 4\,096 = 7.50$$

La variación total es

$$(19)^2 + (16)^2 + (16)^2 + \cdots + (15)^2 + (23)^2 - \frac{(256)}{16} = 4\,244 - 4\,096 = 148.00$$

de modo que la variación que se debe a error resulta

$$148.00 - 16.50 - 6.50 - 111.50 - 7.50 = 6.00$$

Los resultados se muestran en el análisis de varianza de la tabla 9-31. El número total de grados de libertad es $n^2 - 1$ para un cuadrado de $n \times n$. Cada uno de los renglones, las columnas, las letras latinas y las griegas tienen $n - 1$ grados de libertad. Por tanto, los grados de libertad para el error son $n^2 - 1 - 4(n - 1) = (n - 1)(n - 3)$. En este caso, $n = 4$.

Tabla 9-31

Variación	Grados de libertad	Media al cuadrado	F
Renglones (automóviles), 16.50	3	5.500	$\frac{5.500}{2.000} = 2.75$
Columnas (conductores), 6.50	3	2.167	$\frac{2.167}{2.000} = 1.08$
Gasolinas (A, B, C, D), 111.50	3	37.167	$\frac{37.167}{2.000} = 18.6$
Carreteras ($\alpha, \beta, \gamma, \delta$), 7.50	3	2.500	$\frac{2.500}{2.000} = 12.5$
Errores, 6.00	3	2.000	
Total, 148.00	15		

Se tienen $F_{0.95,3,3} = 9.28$ y $F_{0.99,3,3} = 29.5$. En consecuencia, se puede rechazar la hipótesis de que las gasolinas son iguales a un nivel de 0.05, pero no a uno de 0.1.

PROBLEMAS DIVERSOS

9.18. Demostrar que $\sum\alpha_j = 0$ [(15), página 316].

Las medias de la población del tratamiento están dadas por $\mu_j = \mu + \alpha_j$. Se tiene

$$\sum_{j=1}^{a}\mu_j = \sum_{j=1}^{a}\mu + \sum_{j=1}^{a}\alpha_j = a\mu + \sum_{j=1}^{a}\alpha_j = \sum_{j=1}^{a}\mu_j + \sum_{j=1}^{a}\alpha_j$$

donde se ha usado la definición $\mu = (\sum\mu_j)/a$. Se deduce que $\sum\alpha_j = 0$.

9.19. Deducir *a*) ecuación (17), y *b*) ecuación (16) de la página 316.

a) Por definición se tiene

$$\begin{aligned} V_w &= \sum_{j,k}(X_{jk} - \bar{X}_{j\cdot})^2 \\ &= b\sum_{j=1}^{a}\left[\frac{1}{b}\sum_{k=1}^{b}(X_{jk} - \bar{X}_{j\cdot})^2\right] \\ &= b\sum_{j=1}^{a}S_j^2 \end{aligned}$$

donde S_j^2 es la varianza muestral para el tratamiento j-ésimo, como se define en (15), capítulo 5. Entonces, puesto que el tamaño de la muestra es b,

$$\begin{aligned} E(V_w) &= b\sum_{j=1}^{a}E(S_j^2) \\ &= b\sum_{j=1}^{a}\left(\frac{b-1}{b}\sigma^2\right) \\ &= a(b-1)\sigma^2 \end{aligned}$$

utilizando (16) del capítulo 5.

b) Por definición,

$$\begin{aligned} V_b &= b\sum_{j=1}^{a}(\bar{X}_{j\cdot} - \bar{X})^2 \\ &= b\sum_{j=1}^{a}\bar{X}_{j\cdot}^2 - 2b\bar{X}\sum_{j=1}^{a}\bar{X}_{j\cdot} + ab\bar{X}^2 \\ &= b\sum_{j=1}^{a}\bar{X}_{j\cdot}^2 - ab\bar{X}^2 \end{aligned}$$

puesto que

$$\bar{X} = \frac{\sum_j \bar{X}_{j\cdot}^2}{a}$$

Entonces, al omitir el índice de la sumatoria, se tiene

$$E(V_b) = b\sum E(\bar{X}_{j\cdot}^2) - abE(\bar{X})^2 \tag{1}$$

Ahora, para cualquier variable aleatoria U, $E(U^2) = \text{Var}(U) + [E(U)]^2$. De tal modo que

$$E(\bar{X}_{j\cdot}^2) = \text{Var}(\bar{X}_{j\cdot}) + [E(\bar{X}_{j\cdot})]^2 \tag{2}$$

$$E(\bar{X}^2) = \text{Var}(\bar{X}) + [E(\bar{X})]^2 \tag{3}$$

Pero, puesto que las poblaciones del tratamiento son normales, con medias μ_j y varianza común σ^2, del teorema 5-4, página 156, se tiene

$$\text{Var}(\bar{X}_{j\cdot}) = \frac{\sigma^2}{b} \tag{4}$$

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{ab} \tag{5}$$

$$E(\bar{X}_{j\cdot}) = \mu_j = \mu + \alpha_j \tag{6}$$

$$E(\bar{X}) = \mu \tag{7}$$

Usando los resultados de (2) a (7), más el resultado del problema 9.18, en (1), se encuentra que

$$\begin{aligned} E(V_b) &= b\sum\left[\frac{\sigma^2}{b} + (\mu + \alpha_j)^2\right] - ab\left[\frac{\sigma^2}{ab} + \mu^2\right] \\ &= a\sigma^2 + b\sum(\mu + \alpha_j)^2 - \sigma^2 - ab\mu^2 \\ &= (a-1)\sigma^2 + ab\mu^2 + 2b\mu\sum\alpha_j + b\sum\alpha_j^2 - ab\mu^2 \\ &= (a-1)\sigma^2 + b\sum\alpha_j^2 \end{aligned}$$

9.20. Demostrar el teorema 9-1, página 317.

Como se muestra en el problema 9.19*a*),

$$V_w = b\sum_{j=1}^{a} S_j^2 \qquad \text{o} \qquad \frac{V_w}{\sigma^2} = \sum_{j=1}^{a}\frac{bS_j^2}{\sigma^2}$$

donde S_j^2 es la varianza muestral para muestras de tamaño b tomadas de la población del tratamiento j. Mediante el teorema 5-6, página 158, bS_j^2/σ^2 tiene la distribución ji cuadrada con $b - 1$ grados de libertad. Entonces, puesto que las varianzas S_j^2 son independientes, del teorema 4-4, página 121, se concluye que V_w/σ^2 tiene distribución ji cuadrada con $a(b - 1)$ grados de libertad.

9.21. En el problema 9.13 se supone que no hubo diferencias significativas en las replicaciones, esto es, los diferentes días de la semana. ¿Es posible confirmar esta conclusión a un nivel de significancia de *a*) 0.05, *b*) 0.01?

Si hay alguna variación debida a las replicaciones, se incluye en lo que se llama el "residual" o "error aleatorio", $v_e = 84.8$ en la tabla 9-23. Para calcular la variación que se debe a la replicación, se usan los totales de columnas de la tabla 9-21, y se obtiene

$$\begin{aligned} v_{\text{rep}} &= \frac{(57)^2}{8} + \frac{(51)^2}{8} + \frac{(54)^2}{8} + \frac{(47)^2}{8} + \frac{(59)^2}{8} - \frac{(268)^2}{40} \\ &= 1\,807 - 1\,795.6 = 11.4 \end{aligned}$$

Puesto que hay 5 replicaciones, el número de grados de libertad asociado con esta variación es $5 - 1 = 4$. La variación residual después de restar la variación debida a la replicación es $v'_e = 84.8 - 11.4 = 73.4$. Las otras variaciones son iguales a las de la tabla 9-23. La tabla 9-32 es la tabla final del análisis de varianza, tomando en cuenta las replicaciones.

A partir de la tabla se ve que la F calculada para replicación es 1.09. Pero, puesto que $F_{0.95} = 2.71$ para 4, 28 grados de libertad, se concluye que no existe una variación significativa debida a las replicaciones, esto es, los días de la semana no son significativos, a un nivel de 0.05 (y, por tanto, tampoco al de 0.01). Las conclusiones sobre las *Máquinas* y los *Turnos* son iguales a las que se obtuvieron en el problema 9.13.

Tabla 9-32

Variación	Grados de libertad	Cuadrado medio	F
Renglones (máquinas), $v_r = 51.0$	3	17.0	$\frac{17.0}{2.621} = 6.49$
Columnas (turnos), $v_c = 8.1$	1	8.1	$\frac{8.1}{2.621} = 3.05$
Replicaciones (días de la semana), $v_{\text{rep}} = 11.4$	4	2.85	$\frac{2.85}{2.621} = 1.09$
Interacción, $v_i = 6.5$	3	2.167	$\frac{2.167}{2.621} = 0.827$
Aleatoria o residual, $v'_e = 73.4$	28	2.621	
Total, $v = 150.4$	39		

9.22. Describir cómo se pueden usar las técnicas de análisis de varianza para la clasificación tridireccional o los experimentos de tres factores (con entradas únicas). Construir la tabla de análisis de varianza que empleará.

Se supone que la clasificación se divide en A grupos, denotados por $A_1, \ldots, A_\alpha$, B grupos denotados por $B_1, \ldots, B_b$ y C grupos que se denotan por $C_1, \ldots, C_c$. El valor en A_j, B_k y C_l se indica mediante x_{jkl}. El valor $\bar{x}_{jkl.}$, por ejemplo, denota la media de valores en la clase C cuando A_j y B_k se mantiene fijo. Se dan significados similares a $\bar{x}_{j.l}$ y $\bar{x}_{.kl}$. El valor $\bar{x}_{j..}$ es la media de los valores para las clases B y C cuando A_j se mantiene fijo. Finalmente, $\bar{x}$ denota la gran media.

Una *variación total* está dada por

$$v = \sum_{j,k,l} (x_{jkl} - \bar{x})^2 \tag{1}$$

que puede dividirse en siete variaciones, como se indica la tabla 9-33. Estas variaciones son entre clases del mismo tipo y entre clases de diferentes tipos (*interacciones*). La interacción entre todas las clases se denomina *variación residual* o *aleatoria*, como se vio anteriormente.

Las siete variaciones en las cuales se puede dividir (1) son

$$v = v_A + v_B + v_C + v_{AB} + v_{BC} + v_{CA} + v_{ABC}$$

donde

$$v_A = bc \sum_j (\bar{x}_{j..} - \bar{x})^2, \qquad v_B = ca \sum_k (\bar{x}_{.k.} - \bar{x})^2, \qquad v_C = ab \sum_l (\bar{x}_{..l} - \bar{x})^2$$

$$v_{AB} = c \sum_{j,k} (\bar{x}_{jk.} - \bar{x}_{j..} - \bar{x}_{.k.} + \bar{x})^2$$

$$v_{BC} = a \sum_{k,l} (\bar{x}_{.kl} - \bar{x}_{.k} - \bar{x}_{..l} + \bar{x})^2$$

$$v_{CA} = b \sum_{j,l} (\bar{x}_{j.l} - \bar{x}_{..l} - \bar{x}_{j..} + \bar{x})^2$$

$$v_{ABC} = \sum_{j,k,l} (x_{jkl} - \bar{x}_{jk.} - \bar{x}_{j.l} - \bar{x}_{.kl} + \bar{x}_{j..} + \bar{x}_{.k.} + \bar{x}_{..l} - \bar{x})^2$$

Tabla 9-33

Variación	Grados de libertad	Cuadrado medio	F
v_A (Entre los grupos A)	$a - 1$	$\hat{s}_A^2 = \dfrac{v_A}{a-1}$	$\hat{s}_A^2/\hat{s}_{ABC}^2$ $a - 1$, $(a-1)(b-1)(c-1)$ gl
v_B (Entre los grupos B)	$b - 1$	$\hat{s}_B^2 = \dfrac{v_B}{b-1}$	$\hat{s}_B^2/\hat{s}_{ABC}^2$ $b - 1$, $(a-1)(b-1)(c-1)$ gl
v_c (Entre los grupos C)	$c - 1$	$\hat{s}_C^2 = \dfrac{v_C}{c-1}$	$\hat{s}_C^2/\hat{s}_{ABC}^2$ $c - 1$, $(a-1)(b-1)(c-1)$ gl
v_{AB} (Entre los grupos A y B)	$(a-1)(b-1)$	$\hat{s}_{AB}^2 = \dfrac{v_{AB}}{(a-1)(b-1)}$	$\hat{s}_{AB}^2/\hat{s}_{ABC}^2$ $(a-1)(b-1)$, $(a-1)(b-1)(c-1)$ gl
v_{BC} (Entre los grupos B y C)	$(b-1)(c-1)$	$\hat{s}_{BC}^2 = \dfrac{v_{BC}}{(b-1)(c-1)}$	$\hat{s}_{BC}^2/\hat{s}_{ABC}^2$ $(b-1)(c-1)$, $(a-1)(b-1)(c-1)$ gl
v_{CA} (Entre los grupos C y A)	$(c-1)(a-1)$	$\hat{s}_{CA}^2 = \dfrac{v_{CA}}{(c-1)(a-1)}$	$\hat{s}_{CA}^2/\hat{s}_{ABC}^2$ $(c-1)(a-1)$, $(a-1)(b-1)(c-1)$ gl
v_{ABC} (Entre los grupos A, B y C)	$(a-1)(b-1)(c-1)$	$\hat{s}_{ABC}^2 = \dfrac{v_{ABC}}{(a-1)(b-1)(c-1)}$	
v (Total)	$abc - 1$		

PROBLEMAS SUPLEMENTARIOS

CLASIFICACIÓN UNIDIRECCIONAL O EXPERIMENTOS DE UN FACTOR

9.23. Un experimento se ejecuta para determinar las producciones de 5 diferentes variedades de trigo, A, B, C, D, E. Se asignan cuatro lotes de terreno a cada variedad, y las producciones (en quintales por acre) se muestran en la tabla 9-34. Suponiendo que los lotes tienen fertilidad similar y que las variedades se asignan a los lotes de manera aleatoria, determine si existe una diferencia significativa entre las producciones a un nivel de significancia de *a*) 0.05, *b*) 0.01.

Tabla 9-34

A	20	12	15	19
B	17	14	12	15
C	23	16	18	14
D	15	17	20	12
E	21	14	17	18

Tabla 9-35

A	33	38	36	40	31	35
B	32	40	42	38	30	34
C	31	37	35	33	34	30
D	29	34	32	30	33	31

9.24. Una compañía desea probar 4 tipos diferentes de neumáticos, *A*, *B*, *C*, *D*. Los tiempos de vida de los neumáticos, según se determinan a partir de sus dibujos, están dados (en miles de millas) en la tabla 9-35, donde cada tipo ha sido probado en 5 automóviles similares asignados a los neumáticos de manera aleatoria. Pruebe a un nivel de *a*) 0.05, *b*) 0.01, si hay una diferencia entre los neumáticos.

9.25. Un maestro desea probar tres métodos diferentes de enseñar, I, II, III. Para hacerlo, forma tres grupos conformados por 5 estudiantes, cada uno se elige de manera aleatoria, y a cada grupo se le enseña con un método distinto. Después, se aplica el mismo examen a todos los estudiantes, y se obtienen las calificaciones que se presentan en la tabla 9-36. Determine a un nivel de *a*) 0.05, *b*) 0.01 si hay una diferencia significativa entre los métodos pedagógicos.

MODIFICACIONES PARA NÚMEROS DESIGUALES DE OBSERVACIONES

9.26. La tabla 9-37 proporciona el rendimiento en millas por galón que se obtuvieron para automóviles parecidos utilizando 5 marcas diferentes de gasolina. Pruebe a un nivel de significancia *a*) 0.05, *b*) 0.01, si existe una diferencia significativa entre las marcas.

CLASIFICACIÓN BIDIRECCIONAL O EXPERIMENTOS DE DOS FACTORES

9.28. Tres operadores producen los artículos que fabrica una compañía usando 3 máquinas diferentes. El fabricante desea determinar si existe una diferencia entre *a*) operadores, *b*) máquinas. Se ejecuta un experimento para determinar el número de artículos por día que produjo cada operador utilizando cada una de las máquinas. Los resultados se presentan en la tabla 9-39. Proporcione la información que se busca a un nivel de significancia de 0.05.

Tabla 9-39

	Operador 1	Operador 2	Operador 3
Máquina *A*	23	27	24
Máquina *B*	34	30	28
Máquina *C*	28	25	27

9.29. Resuelva el problema 9.28 a un nivel de significancia de 0.01.

9.30. Semillas de 4 diferentes tipos de maíz se siembran en 5 bloques. Cada bloque se divide en 4 lotes, los cuales se asignan de manera aleatoria a los 4 tipos. Pruebe, a un nivel de 0.05 si las producciones en quintales por acre, como se muestra en la tabla 9-40, varían de manera significativa con *a*) tierras diferentes (esto es, los 5 bloques), *b*) diferentes tipos de maíz.

Tabla 9-40

		TIPOS DE MAÍZ			
		I	*II*	*III*	*IV*
BLOQUES	*A*	12	15	10	14
	B	15	19	12	11
	C	14	18	15	12
	D	11	16	12	16
	E	16	17	11	14

9.31. Resuelva el problema 9.30 a un nivel de significancia de 0.01.

9.32. Suponga que en el problema 9.24 la primera observación para cada tipo de neumático se hace con un tipo de automóvil particular, la segunda observación otro tipo particular, etc. Pruebe a un nivel de 0.05 si hay una diferencia entre *a*) los tipos de neumáticos, *b*) los tipos de automóviles.

9.33. Resuelva el problema 9.32 a un nivel de significancia de 0.01.

9.34. En el problema 9.25 la primera entrada para cada método pedagógico corresponde a un estudiante de una escuela específica, la segunda a un estudiante de otra escuela, etc. Pruebe la hipótesis a un nivel de significancia de 0.05, de que existe una diferencia entre *a*) los métodos de enseñanza, *b*) las escuelas.

9.35. Un experimento se lleva a cabo para probar si el color del pelo y las estaturas de las estudiantes en Estados Unidos tienen algún efecto sobre el desempeño académico. Los resultados se dan en la tabla 9-41, donde los números indican 10% de los mejores individuos de los que se van a graduar. Analice el experimento a un nivel de 0.05.

Tabla 9-41

	Pelirroja	Rubia	Morena
Alta	75	78	80
Media	81	76	79
Baja	73	75	77

9.36. Resuelva el problema anterior a un nivel de 0.01.

EXPERIMENTO DE DOS FACTORES CON REPLICACIÓN

9.37. Suponga que el experimento del problema 9.23 se llevó a cabo en el sur de Estados Unidos y que las columnas de la tabla 9-34 indican ahora 4 tipos diferentes de fertilizantes, mientras que en un experimento parecido que se ejecuta en la parte occidental de ese país se producen los resultados de la tabla 9-42. Pruebe, a un nivel de 0.05, si hay una diferencia entre *a*) los fertilizantes *b*) las regiones.

Tabla 9-42

A	16	18	20	23
B	15	17	16	19
C	21	19	18	21
D	18	22	21	23
E	17	18	24	20

9.38. Resuelva el problema 9.37 a un nivel de 0.01.

9.39. La tabla 9-43 proporciona el número de artículos que produjeron 4 diferentes operadores que trabajan con dos tipos de máquinas, *I* y *II*, en distintos días de la semana. Determine, a un nivel de 0.05, si hay diferencias significativas entre *a*) los operadores, *b*) las máquinas.

Tabla 9-43

	Máquina *I*					Máquina *II*				
	Lun	Mar	Miér	Jue	Vie	Lun	Mar	Miér	Jue	Vie
Operador *A*	15	18	17	20	12	14	16	18	17	15
Operador *B*	12	16	14	18	11	11	15	12	16	12
Operador *C*	14	17	18	16	13	12	14	16	14	11
Operador *D*	19	16	21	23	18	17	15	18	20	17

CUADRADO LATINO

9.40. Se realiza un experimento para probar el efecto en la producción de maíz de 4 diferentes fertilizantes, *A*, *B*, *C*, *D* y de las variaciones en la tierra en dos direcciones perpendiculares. Se obtiene el cuadrado latino de la tabla 9-44, donde los números indican la producción de maíz por área unitaria. Pruebe la hipótesis de que no existe una diferencia entre *a*) los fertilizantes y *b*) las variaciones de tierra a un nivel de 0.01.

Tabla 9-44

C 8	*A* 10	*D* 12	*B* 11
A 14	*C* 12	*B* 11	*D* 15
D 10	*B* 14	*C* 16	*A* 10
B 7	*D* 16	*A* 14	*C* 12

9.41. Resuelva el problema 9.40 a un nivel de 0.05.

9.42. Con referencia al problema 9.35, suponga que se introduce un factor adicional que indique la sección *E*, *M* o *W* de Estados Unidos donde nació una estudiante, como se muestra en la tabla 9-45. Determine, a un nivel de 0.05, si existe una diferencia significativa entre el desempeño académico de las estudiantes debido a diferencias en *a*) la estatura, *b*) el color del pelo, *c*) el lugar de nacimiento.

CUADRADOS GRECOLATINOS

9.43. Para producir un tipo de alimento para pollos, 4 cantidades diferentes de cada uno de dos químicos se agregan a los ingredientes básicos. La diferentes cantidades del primer químico están indicadas por A, B, C, D, mientras que las del segundo químico están denotadas por α, β, γ, δ. Se da el alimento a pollitos ordenados en grupos de acuerdo con 4 pesos iniciales distintos, W_1, W_2, W_3, W_4, y cuatro especies, S_1, S_2, S_3, S_4. Los aumentos en peso por tiempo unitario están dados en el cuadrado grecolatino de la tabla 9-46. Aplique el análisis de varianza al experimento a un nivel de significancia de 0.05, e indique cualquier conclusión a la que se pueda llegar.

Tabla 9-46

	W_1	W_2	W_3	W_4
S_1	C_γ 8	B_β 6	A_α 5	D_δ 6
S_2	A_δ 4	D_α 3	C_β 7	B_γ 3
S_3	D_β 5	A_γ 6	B_δ 5	C_α 6
S_4	B_α 6	C_δ 10	D_γ 10	A_β 8

9.44. Cuatro compañías T_1, T_2, T_3, T_4, fabrican cuatro tipos de cables diferentes, C_1, C_2, C_3, C_4. Cuatro operadores, A, B, C, D, utilizando cuatro máquinas α, β, γ, δ, miden las resistencias de los cables. Las resistencias promedio que se obtuvieron se presentan en el cuadrado grecolatino de la tabla 9-47. Lleve a cabo un análisis de varianza a un nivel de 0.05, e indique las conclusiones a las que se pueda llegar.

Tabla 9-47

	C_1	C_2	C_3	C_4
T_1	A_β 164	B_γ 181	C_α 193	D_δ 160
T_2	C_δ 171	D_α 162	A_γ 183	B_β 145
T_3	D_γ 198	C_β 221	B_δ 207	A_α 188
T_4	B_α 157	A_δ 172	D_β 166	C_γ 136

PROBLEMAS DIVERSOS

9.45. La tabla 9-48 proporciona datos sobre la oxidación acumulada en hierro tratado con el químico A, B o C. Determine, a un nivel de significancia de *a*) 0.05 y *b*) 0.01 si hay una diferencia significativa entre los tratamientos.

Tabla 9-48

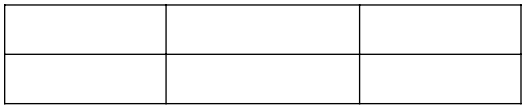
A	3	5	4	4
B	4	2	3	3
C	6	4	5	5

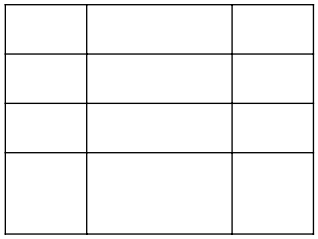
Tabla 9-49

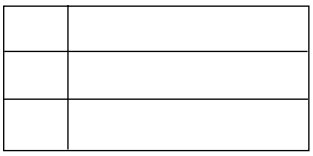
Alto	110	105	118	112	90	
Bajo	95	103	115	107		
Medio	108	112	93	104	96	102

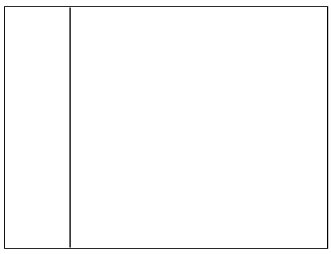
9.46. Un experimento mide los CI de los estudiantes adultos de estatura alta, mediana y baja. Los resultados se presentan en la tabla 9-49. Determine, a un nivel de significancia de *a*) 0.05 y *b*) 0.01 si hay alguna diferencia significativa entre los puntajes de CI relacionados con las diferentes estaturas.

9.47. Se aplica un examen para determinar si veteranos o novatos de diferentes CI tienen mejor desempeño. Las calificaciones que obtuvieron se muestran en la tabla 9-50. Determine, a un nivel de 0.05 si hay una diferencia entre las calificaciones debida a diferencias en *a*) ser veterano o no, *b*) el CI.









9.55. Efectúe un análisis de varianza sobre el cuadrado grecolatino de la tabla 9-55 a un nivel de 0.05, y saque conclusiones.

Tabla 9-54

FACTOR 2 \ FACTOR 1			
	B 16	C 21	A 15
	A 18	B 23	C 14
	C 15	A 18	B 12

Tabla 9-55

FACTOR 2 \ FACTOR 1			
	A_γ 6	B_β 12	C_δ 4	D_α 18
	B_δ 3	A_α 8	D_γ 15	C_β 14
	D_β 15	C_γ 20	B_α 9	A_δ 5
	C_α 16	D_δ 6	A_β 17	B_γ 7

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

9.23. Existe una diferencia significativa en la producción a los dos niveles.

9.24. No hay una diferencia significativa entre los neumáticos a los dos niveles.

9.25. Sí hay una diferencia significativa entre los métodos pedagógicos al nivel de 0.05, pero no al de 0.01.

9.26. Hay una diferencia significativa entre las marcas a un nivel de 0.05, pero no a uno de 0.01.

9.27. Existe una diferencia entre las calificaciones a los dos niveles.

9.28. No existe una diferencia significativa entre los operadores o las máquinas.

9.29. No hay una diferencia significativa entre los operadores o las máquinas.

9.30. Hay una diferencia significativa entre los tipos de maíz, pero no entre las tierras a un nivel de 0.05.

9.31. No existe una diferencia significativa entre el tipo de maíz o la tierra a un nivel de 0.01.

9.32. Hay una diferencia significativa tanto en neumáticos como en automóviles a un nivel de 0.05.

9.33. No existe una diferencia significativa entre neumáticos o automóviles a un nivel de 0.01.

9.34. Hay una diferencia significativa en métodos de enseñanza, pero no entre escuelas a un nivel de 0.05.

9.35. No hay una diferencia significativa tanto en color de pelo como en estatura.

9.36. La misma respuesta que la del problema 9.35.

9.37. Sí hay una diferencia significativa entre las regiones a un nivel de 0.05, pero no entre fertilizantes.

9.38. No existe una diferencia significativa entre las regiones de los fertilizantes a un nivel de 0.01.

9.39. Hay una diferencia significativa entre los operadores, pero no entre las máquinas.

9.40. No hay una diferencia significativa en los fertilizantes o las tierras.

9.41. La misma respuesta que la del problema 9.40.

9.42. No existe una diferencia significativa entre desempeño académico debido a las diferencias en estatura, color de pelo o lugar de nacimiento.

9.43. Existe una diferencia significativa entre las especies y las cantidades del primer químico, pero no existen otras diferencias significativas.

9.44. Hay diferencias significativas entre los tipos de cables, pero no entre las resistencias de los cables debidas a los operadores, las máquinas o las compañías.

9.45. No existe una diferencia significativa en tratamientos a cualquier nivel.

9.46. No hay una diferencia significativa en los puntajes de CI a cualquier nivel.

9.47. Hay diferencias significativas entre calificaciones de exámenes debidas tanto al estatus de ser veterano como al CI a un nivel de 0.05.

9.48. A un nivel de 0.01, las diferencias entre las calificaciones de exámenes debidas al estatus de ser veterano no son significativas, pero las que se deben al CI sí lo son.

9.49. No hay diferencias significativas entre calificaciones de exámenes de estudiantes de diferentes partes del país, pero sí las hay entre calificaciones debidas al CI.

9.50. La misma respuesta que la del problema 9.49.

9.51. Hay una diferencia significativa debida a los químicos o regiones a un nivel de 0.05.

9.52. Existen diferencias significativas que se deben a las regiones, pero no a los fertilizantes.

9.53. No existen diferencias significativas debidas a las regiones o a los fertilizantes.

9.54. No hay diferencias significativas debidas al factor 1, al factor 2 o a los tratamientos *A*, *B*, *C*.

9.55. No existen diferencias significativas debidas a los factores o a los tratamientos.

PROBLEMAS APORTADOS

1. Un fabricante a pequeña escala de latas, con apenas 5 productos a enlatar, necesita evaluar la resistencia a la corrosión de cuatro tipos de aleaciones. Considere que cada uno de los productos son variados en contenido y tamaños, tales como aceite vegetal, lubricante de motor, néctares, chiles y frutas en almíbar. Es apropiado un diseño de bloques completos aleatorizados para controlar el efecto de esta variable. Cada lata constituye un bloque, y cada tipo de aleación un tratamiento. Se coloca una aleación en cada tipo de producto que habrá de enlatarse. Luego se ponen en cuarentena, tras la cual se mide la resistencia a la corrosión de la aleación en mm/año. Los datos obtenidos y las estadísticas de resumen pertinentes se muestran en la siguiente tabla. Calcular:

a) la media de las producciones para los diferentes tratamientos,
b) la gran media para todos los tratamientos,
c) la variación total,
d) la variación entre tratamientos,
e) la variación dentro de tratamientos.

Utilizar el método largo.

Tratamientos (tipo de aleación)

Bloques (productos a enlatar)	1	2	3	4
1	9.1	17.1	20.8	11.8
2	13.4	20.3	28.3	16.0
3	15.6	24.6	23.7	16.2
4	11.0	18.2	21.4	14.1
5	12.7	19.8	25.1	15.8

2. Realizar la estimación insesgada de la varianza poblacional σ^2 a partir de *a*) la variación entre tratamientos con base en la hipótesis nula de medias de tratamientos iguales, y *b*) la varianza dentro de tratamientos.

CAPÍTULO 10

INTRODUCCIÓN

La mayoría de las pruebas de hipótesis y significancia (o reglas de decisión) que se consideraron en los capítulos anteriores, requieren de diversos supuestos acerca de la distribución de la población de la cual se toman las muestras. Por ejemplo, en el capítulo 5, a menudo se requiere que las distribuciones poblacionales sean normales o casi normales.

En la práctica, surgen situaciones en las que esas suposiciones no se pueden justificar pues existe duda en cuanto a su aplicación, como el caso en que una población sea muy sesgada. Por tanto, los estadísticos han elaborado diversas pruebas y métodos que son independientes de las distribuciones poblacionales y los parámetros asociados, las cuales se llaman *pruebas no paramétricas*.

Las pruebas no paramétricas se pueden usar como sustituciones abreviadas de pruebas más complicadas. Son especialmente valiosas para manejar datos no numéricos, como los que surgen cuando consumidores clasifican cereales u otros productos en orden de preferencia.

PRUEBA DE LOS SIGNOS

Considere la tabla 10-1, que muestra los números de pernos defectuosos producidos por dos diferentes tipos de máquinas (I y II) durante 12 días consecutivos, y suponga que las máquinas tienen la misma producción total por día. Se desea probar la hipótesis H_0 de que no existe una diferencia entre las máquinas: que las diferencias observadas entre las máquinas en términos de los números de pernos defectuosos que producen sólo se debe a la casualidad, es decir, que las muestras se toman de la misma población.

Una prueba no paramétrica sencilla en el caso de muestras emparejadas como las anteriores la proporciona la *prueba de los signos*. Esta prueba consiste en tomar la diferencia entre los números de pernos defectuosos de cada día y escribir sólo los *signos* de la diferencia; por ejemplo, para el día 1 se tiene $47 - 71$, la cual es negativa. De este modo se obtiene la secuencia de signos de la tabla 10-1

$$- \quad - \quad + \quad - \quad - \quad - \quad + \quad - \quad + \quad - \quad - \quad - \tag{1}$$

(esto es, 3 signos más y 9 signos menos). Ahora, si es igualmente probable obtener un $+$ como un $-$, se esperaría obtener 6 de cada uno. La prueba de H_0 es, en consecuencia, equivalente a la de si en 12 lanzamientos de una moneda legal resultan 3 caras $(+)$ y 9 cruces $(-)$. Lo anterior implica la distribución binomial del capítulo 4. El problema 10.1 muestra que al usar una prueba de dos colas de esta distribución a un nivel de significancia de 0.05, no es posible rechazar H_0; esto es, no hay una diferencia entre las máquinas a este nivel.

Nota 1 Si durante algún día las máquinas producen el mismo número de pernos defectuosos, una diferencia de *cero* aparecería en la secuencia (1), en cuyo caso se pueden omitir estos valores muestrales y usar 11 en vez de 12 observaciones.

Nota 2 También es posible usar una aproximación normal para la distribución binomial mediante una corrección de continuidad (vea el problema 10.2).

Tabla 10-1

Día	1	2	3	4	5	6	7	8	9	10	11	12
Máquina I	47	56	54	49	36	48	51	38	61	49	56	52
Máquina II	71	63	45	64	50	55	42	46	53	57	75	60

Aunque la prueba de los signos es particularmente útil para las muestras emparejadas, como en la tabla 10-1, se puede usar también para los problemas que contienen muestras simples (vea los problemas 10.3 y 10.4).

PRUEBA *U* DE MANN-WHITNEY

Considere la tabla 10-2, la cual muestra las resistencias de cables que se hicieron de dos aleaciones diferentes, I y II. En esta tabla se tienen dos muestras: 8 cables de la aleación I y 10 de la II. Se desea decidir si hay una diferencia o no entre las muestras o, de manera equivalente, si son de la misma muestra poblacional o no. Aunque este problema se puede resolver por medio de la prueba *t* del capítulo 7, es útil una prueba no paramétrica, que se denomina *prueba* U *de Mann-Whitney*, o, de forma breve, *prueba* U. Esta prueba consta de los siguientes pasos:

Tabla 10-2

Aleación I				Aleación II				
18.3	16.4	22.7	17.8	12.6	14.1	20.5	10.7	15.9
18.9	25.3	16.1	24.2	19.6	12.9	15.2	11.8	14.7

Paso 1. Combine todos los valores muestrales en un arreglo del más pequeño al más grande, y asigne rangos (en este caso de 1 a 18) a todos estos valores. Si dos o más valores muestrales son idénticos (esto es, existen *puntajes empatados*, o simplemente, *empates*), a cada uno de los valores muestrales se le asigna un rango igual a la *media* de los rangos que se asignarían de otra manera. Si la entrada 18.9 de la tabla 10-2 fuera 18.3, dos valores idénticos, 18.3, ocuparían los rangos de 12 y 13 en el arreglo, de modo que el rango que se asigna a cada uno sería $\frac{1}{2}(12 + 13) = 12.5$.

Paso 2. Encuentre la suma de los rangos para cada una de las muestras y denotarlas como R_1 y R_2, donde N_1 y N_2 son los tamaños respectivos de las muestras. Por conveniencia, se elige N_1 como el tamaño más pequeño si son desiguales, por lo que $N_1 \leq N_2$. Una diferencia significativa entre el rango de las sumas R_1 y R_2 implica una diferencia significativa entre las muestras.

Paso 3. Para probar la diferencia entre el rango de las sumas utilizar el estadístico

$$U = N_1N_2 + \frac{N_1(N_1 + 1)}{2} - R_1 \tag{2}$$

que corresponde a la muestra 1. La distribución muestral de U es simétrica, y su media y varianza están dadas, respectivamente, por las fórmulas

$$\mu_U = \frac{N_1N_2}{2} \qquad \sigma_U^2 = \frac{N_1N_2(N_1 + N_2 + 1)}{12} \tag{3}$$

Si $N_1 + N_2$ es grande (mayor a 20), entonces U tiene una distribución casi normal, de modo que

$$Z = \frac{U - \mu_U}{\sigma_U} \tag{4}$$

está distribuida normalmente con media 0 y varianza 1. Utilizando el apéndice C se puede decidir si las muestras son diferentes o no de manera significativa. El problema 10.5 muestra que existe una diferencia significativa entre los cables a un nivel de 0.05.

Nota 3. El valor correspondiente a la muestra 2, que está dado por el estadístico

$$U = N_1N_2 + \frac{N_2(N_2 + 1)}{2} - R_2 \tag{5}$$

tiene la misma distribución muestral que el estadístico (2), con la media y la varianza de las fórmulas (3). El estadístico (5) se relaciona con el (2), ya que si U_1 y U_2 son los valores que corresponden a los estadísticos (2) y (5), respectivamente, se llega al resultado

$$U_1 + U_2 = N_1N_2 \tag{6}$$

Se tiene también

$$R_1 + R_2 = \frac{N(N + 1)}{2} \tag{7}$$

donde $N = N_1 + N_2$. El resultado (7) proporciona una verificación de los cálculos.

Nota 4. El estadístico U en la ecuación (2) es el número total de veces en que los valores de la muestra 1 preceden a los de la muestra 2 cuando todos los valores muestrales se ordenan de acuerdo con una magnitud creciente, lo cual proporciona un *método de conteo* alternativo para encontrar U.

PRUEBA *H* DE KRUSKAL-WALLIS

La prueba U es una prueba no paramétrica para decidir si dos muestras son o no de la misma población. Una generalización de ella para k muestras la proporciona la *prueba* H *de Kruskal-Wallis* o, simplemente, *prueba* H.

Es posible describir la prueba H de la siguiente manera: suponga que se tienen k muestras de tamaños $N_1, N_2, \ldots, N_k$ y con el tamaño total de todas las muestras juntas dado por $N = N_1 + N_2 + \cdots + N_k$. Además, suponga que se clasifican los datos de todas las muestras juntas y que las sumas de los rangos para las k muestras son $R_1, R_2, \ldots, R_k$, respectivamente. Si se define el estadístico como

$$H = \frac{12}{N(N + 1)}\sum_{j=1}^{k}\frac{R_j^2}{N_j} - 3(N + 1) \tag{8}$$

entonces es posible demostrar que la distribución muestral de H es casi una *distribución ji-cuadrada* con $k - 1$ grados de libertad, siempre que $N_1, N_2, \ldots, N_k$ sean por lo menos 5.

La prueba H proporciona un método no paramétrico en el *análisis de varianza* para clasificación unidireccional, o experimentos de un factor, y se pueden hacer generalizaciones.

PRUEBA *H* CORREGIDA PARA EMPATES

En el caso en que existan demasiados empates entre las observaciones de los datos muestrales, el valor de H dado por el estadístico (8) es menor de lo que debe ser. El valor corregido H, denotado por H_c, se obtiene dividiendo el valor obtenido del estadístico (8) entre el factor de corrección

$$1 - \frac{\sum(T^3 - T)}{N^3 - N} \tag{9}$$

donde T es el número de empates que corresponde a cada observación y donde la suma incluye todas las observaciones. Si no existen empates, entonces $T = 0$ y el factor (9) se reduce a 1, por lo que no se necesita una corrección. En la práctica, por lo general la corrección es insignificante (esto es, no es suficiente para merecer un cambio en la decisión).

PRUEBA DE CORRIDAS (RACHAS) DE ALEATORIEDAD

Aunque la palabra *aleatorio* se ha usado muchas veces en este libro (como en "muestreo aleatorio" y "lanzar una moneda de manera aleatoria"), ningún capítulo ha incluido una prueba de aleatoriedad. La *teoría de corridas (rachas)* proporciona una prueba no paramétrica para lograr lo anterior.

Para entender lo que es una corrida, considere una secuencia formada por dos símbolos, a y b, como

$$a \quad a \;\big|\; b \quad b \quad b \;\big|\; a \;\big|\; b \quad b \;\big|\; a \quad a \quad a \quad a \quad a \;\big|\; b \quad b \quad b \;\big|\; a \quad a \quad a \quad a \;\big|\; \tag{10}$$

Al lanzar una moneda, por ejemplo, *a* podría representar caras y *b* cruces. O en el muestreo de los pernos producidos por una máquina, *a* podría representar defectuosos y *b* no defectuosos.

Una *corrida* se define como un conjunto de símbolos idénticos (o relacionados) que se encuentra contenido entre dos símbolos distintos o ningún símbolo (como al principio o al final de una secuencia). Al proceder de izquierda a derecha en la secuencia (10), la primera corrida, indicada por una barra vertical, consta de dos *a*; asimismo, la segunda corrida consta de tres *b*, la tercera de una *a*, etc. Hay siete corridas en total.

Es evidente que existe alguna relación entre el carácter aleatorio y el número de corridas. Por tanto, para la secuencia

$$a \mid b \mid a \mid b \mid a \mid b \mid a \mid b \mid a \mid b \mid a \mid b \tag{11}$$

hay un *patrón cíclico* en el cual se va de *a* a *b*, se regresa a *a*, etc., lo que sería difícil considerar como aleatorio. En ese caso existen *demasiadas* corridas (de hecho, se tiene el número máximo posible para el número dado de las *a* y las *b*).

Por otro lado, para la secuencia

$$a\ a\ a\ a\ a\ a \mid b\ b\ b\ b \mid a\ a\ a\ a\ a \mid b\ b\ b \mid \tag{12}$$

parece que hay un *patrón de tendencia*, en el cual las *a* y *b* se agrupan (o acumulan), en cuyo caso existen *muy pocas* corridas, y no se podría considerar la secuencia como aleatoria.

Por tanto, una secuencia se consideraría como no aleatoria si contiene demasiadas o insuficientes corridas, y como aleatoria en caso contrario. Para cuantificar la idea anterior, suponga que se forman todas las secuencias posibles que constan de las N_1 *a* y N_2 *b* para un total de N símbolos ($N_1 + N_2 = N$). El conjunto de todas estas secuencias proporciona una distribución muestral. Cada secuencia tiene un número asociado de corridas, denotado por V, de modo que así se llega a la distribución muestral del estadístico V. Es posible demostrar que la distribución muestral anterior tiene una media y una varianza que están dadas, respectivamente, por las fórmulas

$$\mu_V = \frac{2N_1N_2}{N_1 + N_2} + 1 \qquad \sigma_V^2 = \frac{2N_1N_2(2N_1N_2 - N_1 - N_2)}{(N_1 + N_2)^2(N_1 + N_2 - 1)} \tag{13}$$

Mediante la fórmula (13) se puede probar la hipótesis del carácter aleatorio a niveles adecuados de significancia. Resulta que si tanto N_1 como N_2 son equivalentes por lo menos a 8, entonces la distribución muestral de V es casi una distribución normal. En consecuencia,

$$Z = \frac{V - \mu_V}{\sigma_V} \tag{14}$$

está distribuida de manera casi normal estándar con media 0 y varianza 1, por lo que se puede usar el apéndice C.

APLICACIONES ADICIONALES PARA LA PRUEBA DE CORRIDAS

A continuación se describen otras aplicaciones de la prueba de corridas para problemas estadísticos:

1. **PRUEBA ARRIBA Y ABAJO DE LA MEDIANA DEL CARÁCTER ALEATORIO DE DATOS NUMÉRICOS.** Para determinar si los datos numéricos (como los que se obtienen en una muestra) son aleatorios, es necesario colocar primero los datos en *el mismo orden* en el que se obtuvieron. Después, debe encontrarse la mediana de los datos y sustituir cada entrada por una letra, *a* o *b*, según su valor esté *arriba* o *abajo* de la mediana. Si un valor es igual a la mediana, se omite de la muestra. Que la muestra se considere aleatoria o no depende de que la secuencia de las *a* y *b* sea o no aleatoria. (Vea el problema 10.20.)

2. **DIFERENCIAS ENTRE POBLACIONES DE LAS CUALES SE TOMAN MUESTRAS.** Suponga que dos muestras de tamaños m y n se denotan por $a_1, a_2, \ldots, a_m$ y $b_1, b_2, \ldots, b_n$, respectivamente. Para decidir si las muestras son de la misma población o no, primero se deben ordenar todos los valores muestrales $m + n$ en una secuencia de valores crecientes. Si algunos de los valores son idénticos se deben ordenar con base en un proceso aleatorio (como usar números aleatorios). Si la secuencia resultante es aleatoria, se puede concluir que las muestras no son realmente diferentes y, en consecuencia, son de la misma población; si la secuencia no es aleatoria, no es posible llegar a la *conclusión* anterior. Esta prueba puede proporcionar una alternativa a la de U de Mann-Whitney. (Vea el problema 10.21.)

CORRELACIÓN DE RANGOS DE SPEARMAN

Los métodos no paramétricos también se pueden usar para medir la correlación de dos variables, X y Y. En vez de usar los valores precisos de las variables, o cuando dicha precisión no se puede alcanzar, se ordenan los datos de 1 a N de acuerdo con el tamaño, la importancia, etc. Si X y Y se clasifican de tal manera, el *coeficiente de la correlación de rangos*, o *fórmula de Spearman para correlación de rangos* (como con frecuencia se le conoce), está dada por

$$r_S = 1 - \frac{6\sum D^2}{N(N^2 - 1)} \tag{15}$$

donde D denota las diferencias entre los rangos de valores correspondientes de X y Y, y donde N es el número de pares de valores (X, Y) de los datos.

PROBLEMAS RESUELTOS

PRUEBA DE LOS SIGNOS

10.1. Con referencia a la tabla 10-1, probar la hipótesis H_0 de que no existe una diferencia entre las máquinas I y II contra la hipótesis alternativa H_1 de que sí hay una diferencia a un nivel de significancia de 0.05.

La figura 10-1 es una gráfica de la distribución binomial (y una aproximación normal a ella) que proporciona las probabilidades de x caras en 12 lanzamientos de una moneda legal, donde $x = 0, 1, 2, \ldots, 12$. Del capítulo 4, la probabilidad de x caras es

$$\Pr\{x\} = \binom{12}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{12-x} = \binom{12}{x}\left(\frac{1}{2}\right)^{12}$$

por lo que $\Pr\{0\} = 0.00024$, $\Pr\{1\} = 0.00293$, $\Pr\{2\} = 0.01611$ y $\Pr\{3\} = 0.05371$.

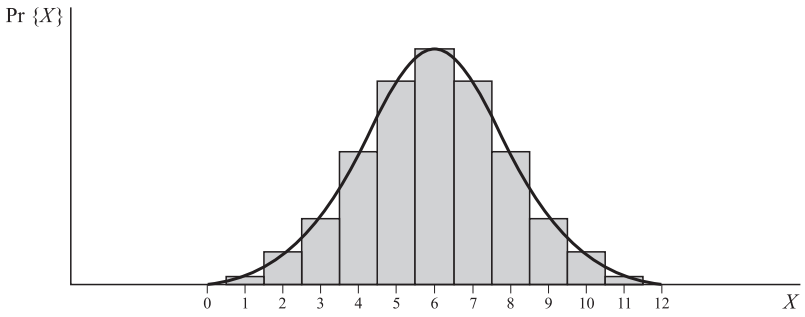


Figura 10-1

Puesto que H_1 es la hipótesis de que existe una *diferencia* entre las máquinas, en vez de que la máquina I es *mejor* que la II, se usa una prueba de dos colas. A un nivel de significancia de 0.05, cada cola tiene la probabilidad asociada $\frac{1}{2}(0.05) = 0.025$. Ahora, se suman las probabilidades en la cola de la izquierda hasta que la suma exceda 0.025. Por tanto,

$$\Pr\{0, 1 \text{ o bien } 2 \text{ caras}\} = 0.00024 + 0.00293 + 0.01611 = 0.01928$$

$$\Pr\{0, 1\ 2 \text{ o bien } 3 \text{ caras}\} = 0.00024 + 0.00293 + 0.01611 + 0.05371 = 0.07299$$

Puesto que 0.025 es mayor que 0.01928 pero menor que 0.07299, es posible rechazar la hipótesis H_0 si el número de caras es 2 o menor (o, mediante simetría, si el número de caras es 10 o mayor); no obstante, el número de las caras (los signos + en la secuencia (1) de este capítulo) es 3. En consecuencia, no se puede rechazar H_0 a un nivel de 0.05, y es necesario concluir que no existe diferencia entre las máquinas a este nivel.

10.2. Resolver el problema 10.1 mediante una aproximación normal a la distribución binomial.

Para una aproximación normal a la distribución binomial se emplea el hecho de que el puntaje z que corresponde al número de caras es

$$Z = \frac{X - \mu}{\sigma} = \frac{X - Np}{\sqrt{Npq}}.$$

Puesto que la variable X de la distribución binomial es discreta mientras que la de una distribución normal es continua, se hace una *corrección para continuidad* (por ejemplo, 3 caras son realmente un valor entre 2.5 y 3.5 caras). Lo anterior equivale a decrecer X por 0.5 si $X > Np$ y a incrementar X por 0.5 si $X < Np$. Ahora, $N = 12$, $\mu = Np = (12)(0.5) = 6$ y $\sigma = \sqrt{Npq} = \sqrt{(12)(0.5)(0.5)} = 1.73$, de modo que

$$z = \frac{(3 + 0.5) - 6}{1.73} = -1.45$$

Puesto que lo anterior es mayor que -1.96 (el valor de z para el cual el área en la cola de la izquierda es 0.025) se llega a la misma conclusión que la del problema 10.1.

Observemos que $\Pr\{Z \leq -1.45\} = 0.0735$, lo que concuerda muy bien con el $\Pr\{X \leq 3 \text{ cabezas}\} = 0.07299$ del problema 10.1.

10.3. La Compañía PQR afirma que el tiempo de vida de un tipo de pila que fabrica es mayor a 250 horas. Un defensor de consumidores que desea determinar si la afirmación se justifica mide los tiempos de vida de 24 de las pilas de la compañía. La tabla 10-3 presenta los resultados. Suponiendo que la muestra es aleatoria, determinar si la afirmación de la compañía está justificada a un nivel de significancia de 0.05.

Tabla 10-3

271	230	198	275	282	225	284	219
253	216	262	288	236	291	253	224
264	295	211	252	294	243	272	268

Sea H_0 la hipótesis de que las pilas de la compañía tienen un tiempo de vida equivalente a 250 horas, y sea H_1 la hipótesis de que tienen un tiempo de vida mayor a dichas horas. Para probar la hipótesis H_0 contra la H_1, es posible usar la prueba de los signos de la siguiente manera: se resta 250 de cada entrada de la tabla 10-3 y se registran los signos de las diferencias como se muestra en la tabla 10-4. Se ve que hay 15 signos más y 8 signos menos.

Tabla 10-4

+	−	−	+	+	−	+	−
+	−	+	+	−	+	+	−
+	+	−	+	+	−	+	+

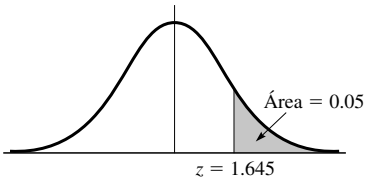


Figura 10-2

Usando la prueba de una cola a un nivel de significancia de 0.05 se rechaza H_0 si los puntajes z fueran mayores que 1.645 (figura 10-2). Puesto que el puntaje z, con base en una corrección de continuidad, es

$$z = \frac{(15 - 0.5) - (24)(0.5)}{\sqrt{(24)(0.5)(0.5)}} = 1.02$$

no es posible justificar la afirmación de la compañía a un nivel de 0.05.

10.4. Una muestra de 40 calificaciones de un examen estatal se muestra en la tabla 10-5. Probar la hipótesis a un nivel de significancia de 0.05 de que la mediana de las calificaciones de todos los participantes es *a*) 66, *b*) 75.

Tabla 10-5

71	67	55	64	82	66	74	58	79	61
78	46	84	93	72	54	78	86	48	52
67	95	70	43	70	73	57	64	60	83
73	40	78	70	64	86	76	62	95	66

a) Restando 66 de todas las entradas en la tabla 10-5, y conservando los signos asociados, resulta la tabla 10-6, donde observamos que hay 23 signos más, 15 signos menos y 2 ceros. Si se descartan los 2 ceros, la muestra consiste en 38 signos: 23 más y 15 menos. Con base en una prueba de dos colas de la distribución normal con probabilidades $\frac{1}{2}(0.05) = 0.025$ en cada cola (figura 10-3), adoptamos la siguiente regla de decisión:

Aceptar la hipótesis si $-1.96 \leq z \leq 1.96$.
De lo contrario, rechazarla.

Tabla 10-6

+	+	−	−	+	0	+	−	+	−
+	−	+	+	+	−	+	+	−	−
+	+	+	−	+	+	−	−	−	+
+	−	+	+	−	+	+	−	+	0

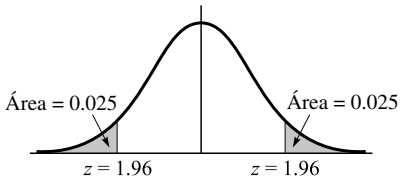


Figura 10-3

Puesto que $$Z = \frac{X - Np}{\sqrt{Npq}} = \frac{(23 - 0.5) - (38)(0.5)}{\sqrt{(38)(0.5)(0.5)}} = 1.14$$

se acepta la hipótesis de que la mediana es 66 a un nivel de 0.05.

Observemos que fue posible también usar 15, el número de los signos menos, en cuyo caso

$$z = \frac{(15 + 0.5) - (38)(0.5)}{\sqrt{(38)(0.5)(0.5)}} = -1.14$$

con la misma conclusión.

b) Restando 75 de todas las entradas de la tabla 10-5 resulta la tabla 10-7, en la que existen 13 signos más y 27 signos menos. Puesto que

$$z = \frac{(13 + 0.5) - (40)(0.5)}{\sqrt{(40)(0.5)(0.5)}} = -2.06$$

rechazamos la hipótesis de que la mediana es 75 a un nivel de 0.05.

Tabla 10-7

−	−	−	−	+	−	−	−	+	−
+	−	+	+	−	−	+	+	−	−
−	+	−	−	−	−	−	−	−	+
−	−	+	−	−	+	+	−	+	−

Usando este método es posible llegar a un intervalo de confianza de 95% para la calificación mediana del examen. (Vea el problema 10.30.)

PRUEBA *U* DE MANN-WHITNEY

10.5. Con referencia a la tabla 10-2 determinar si existe una diferencia a un nivel de significancia de 0.05 entre los cables de la aleación I y los de la II.

Se puede organizar la tarea de acuerdo con los pasos 1, 2 y 3 (descritos antes en este capítulo):

Paso 1. Al combinar los 18 valores muestrales en un arreglo del más pequeño al más grande, se tiene la primera línea de la tabla 10-8. Se asignan a estos valores los números 1 a 18 en la segunda línea, que representa los rangos.

Paso 2. Para encontrar la suma de los rangos para cada muestra se reescribe la tabla 10-2 usando los rangos asociados de la tabla 10-8; el resultado es la tabla 10-9. La suma de los rangos es 106 para el caso de la aleación I y 65 para el de la II.

Tabla 10-8

10.7	11.8	12.6	12.9	14.1	14.7	15.2	15.9	16.1	16.4	17.8	18.3	18.9	19.6	20.5	22.7	24.2	25.3
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Paso 3. Puesto que la muestra de la aleación I tiene el tamaño más pequeño, $N_1 = 8$ y $N_2 = 10$. Las sumas correspondientes de los rangos son $R_1 = 106$ y $R_2 = 65$. Entonces,

$$U = N_1N_2 + \frac{N_1(N_1 + 1)}{2} - R_1 = (8)(10) + \frac{(8)(9)}{2} - 106 = 10$$

$$\mu_U = \frac{N_1N_2}{2} = \frac{(8)(10)}{2} = 40 \qquad \sigma_U^2 = \frac{N_1N_2(N_1 + N_2 + 1)}{12} = \frac{(8)(10)(19)}{12} = 126.67$$

Por tanto, $\sigma_U = 11.25$ y

$$Z = \frac{U - \mu_U}{\sigma_U} = \frac{10 - 40}{11.25} = -2.67$$

Tabla 10-9

Aleación I		Aleación II	
Resistencia del cable	Rango	Resistencia del cable	Rango
18.3	12	12.6	3
16.4	10	14.1	5
22.7	16	20.5	15
17.8	11	10.7	1
18.9	13	15.9	8
25.3	18	19.6	14
16.1	9	12.9	4
24.2	17	15.2	7
Suma	106	11.8	2
		14.7	6
		Suma	65

Puesto que la hipótesis H_0 que se desea probar es si *no* hay una diferencia entre las aleaciones, se requiere una prueba de dos colas. Para el nivel de significancia de 0.05 se tiene la regla de decisión:

Aceptar H_0 si $-1.96 \leq z \leq 1.96$.
De lo contrario, rechazarla.

Puesto que $z = -2.67$, rechazamos H_0 y concluimos que existe una diferencia entre las aleaciones a un nivel de 0.05.

10.6. Verificar los resultados (6) y (7) de este capítulo con los datos del problema 10.5.

a) Puesto que las muestras 1 y 2 producen valores de U dados por

$$U_1 = N_1N_2 + \frac{N_1(N_1 + 1)}{2} - R_1 = (8)(10) + \frac{(8)(9)}{2} - 106 = 10$$

$$U_2 = N_1N_2 + \frac{N_2(N_2 + 1)}{2} - R_2 = (8)(10) + \frac{(10)(11)}{2} - 65 = 70$$

se tiene $U_1 + U_2 = 10 + 70 = 80$ y $N_1N_2 = 106 + 65 = 171$ y

$$\frac{N(N+1)}{2} = \frac{(N_1 + N_2)(N_1 + N_2 + 1)}{2} = \frac{(18)(19)}{2} = 171$$

10.7. Resolver el problema 10.5 utilizando el estadístico U de la muestra de la aleación II.

Para el caso de la muestra de la aleación II,

$$U = N_1N_2 + \frac{N_2(N_2+1)}{2} - R_2 = (8)(10) + \frac{(10)(11)}{2} - 65 = 70$$

por lo que

$$Z = \frac{U - \mu_U}{\sigma_U} = \frac{70 - 40}{11.25} = 2.67$$

Este valor de z es el *negativo* del de z del problema 10.5, y la cola de la derecha de la distribución normal se usa en vez del de la izquierda. Puesto que este valor de z se ubica también fuera de $1.96 \leq z \leq 1.96$, la conclusión es igual a la del problema 10.5.

10.8. Un profesor tiene dos clases de psicología: una de 9 estudiantes en las mañanas y otra de 12 estudiantes en las tardes. En un examen final programado al mismo tiempo para todos los estudiantes, los grupos recibieron las calificaciones que se muestran en la tabla 10-10. ¿Es posible concluir a un nivel de significancia de 0.05 que el grupo matutino tiene mejor desempeño que el vespertino?

Tabla 10-10

Grupo matutino	73	87	79	75	82	66	95	75	70			
Grupo vespertino	86	81	84	88	90	85	84	92	83	91	53	84

Paso 1. La tabla 10-11 muestra el arreglo de calificaciones y rangos. Observe que el rango de las dos calificaciones de 75 es $\frac{1}{2}(5 + 6) = 5.5$, mientras que el rango de las tres calificaciones de 84 es $\frac{1}{3}(11 + 12 + 13) = 12$.

Paso 2. Reescribiendo la tabla 10-10 en términos de los rangos se produce la tabla 10-12.

Verificar: $R_1 = 73$, $R_2 = 158$ y $N = N_1 + N_2 = 9 + 12 = 21$; por tanto, $R_1 + R_2 = 73 + 158 = 231$ y

$$\frac{N(N+1)}{2} = \frac{(21)(22)}{2} = 231 = R_1 + R_2$$

Tabla 10-11

53	66	70	73	75	75	79	81	82	83	84	84	84	85	86	87	88	90	91	92	95
1	2	3	4	5.5		7	8	9	10	12			14	15	16	17	18	19	20	21

Tabla 10-12

													Suma de rangos
Grupo matutino	4	16	7	5.5	9	2	21	5.5	3				73
Grupo vespertino	15	8	12	17	18	14	12	20	10	19	1	12	158

Paso 3.

$$U = N_1N_2 + \frac{N_1(N_1 + 1)}{2} - R_1 = (9)(12) + \frac{(9)(10)}{2} - 73 = 80$$

$$\mu_U = \frac{N_1N_2}{2} = \frac{(9)(12)}{2} = 54 \qquad \sigma_U^2 = \frac{N_1N_2(N_1 + N_2 + 1)}{12} = \frac{(9)(12)(22)}{12} = 198$$

Por tanto,

$$Z = \frac{U - \mu_U}{\sigma_U} = \frac{80 - 54}{14.07} = -1.85$$

Puesto que se desea probar la hipótesis H_1 de que el grupo matutino tiene peor desempeño que el vespertino contra la hipótesis H_0 de que no existe una diferencia a un nivel de 0.05, es necesario llevar a cabo la prueba de una cola. Con referencia a la figura 10-2, que se aplica aquí, se tiene la regla de decisión:

Aceptar H_0 si $z \leq 1.645$.
Rechazar H_0 si $z > 1.645$.

Puesto que el valor verdadero de $z = 1.85 > 1.645$, rechazamos H_0 y se concluye que el grupo de la mañana tuvo peor desempeño que el de la tarde a un nivel de 0.05. Sin embargo, no se puede llegar a la conclusión anterior a un nivel de 0.01 (vea el problema 10.33).

10.9. Calcular U con los datos en la tabla 10-13 usando a) la fórmula (2) de este capítulo, b) el método de conteo (como se describe en la nota 4 de este capítulo).

a) Ordenando los datos de las dos muestras en un arreglo de magnitud creciente y asignando rangos de 1 a 5 resulta la tabla 10-14. Luego, si se sustituyen los datos de la tabla 10-13 con los rangos correspondientes se produce la tabla 10-15, a partir de la cual las sumas de los rangos son $R_1 = 5$ y $R_2 = 10$. Puesto que $N_1 = 2$ y $N_2 = 3$, el valor de U es

$$U = N_1N_2 + \frac{N_1(N_1 + 1)}{2} - R_1 = (2)(3) + \frac{(2)(3)}{2} - 5 = 4$$

De manera similar, encontramos que el valor U para la muestra 2 es $U = 2$.

Tabla 10-13

Muestra 1	22	10	
Muestra 2	17	25	14

Tabla 10-14

Datos	10	14	17	22	25
Rango	1	2	3	4	5

Tabla 10-15

				Suma de rangos
Muestra 1	4	1		5
Muestra 2	3	5	2	10

b) Sustituimos los valores muestrales de la tabla 10-14 con I o II, según el valor pertenezca a la muestra 1 o 2. Entonces, la primera línea de la tabla 10-14 se convierte en

Datos	I	II	II	I	II

De esto observamos que

Número de valores de la muestra 1 que precede al primer valor de la muestra 2 = 1
Número de valores de la muestra 1 que precede al segundo valor de la muestra 2 = 1
Número de valores de la muestra 1 que precede al tercer valor de la muestra 2 = 2
Total = 4

Por tanto, el valor U que corresponde a la primera muestra es 4.

De manera similar, se tiene

Número de valores de la muestra 2 que precede al primer valor de la muestra 1 = 0
Número de valores de la muestra 2 que precede al segundo valor de la muestra 1 = 2
Total = 2

Entonces, el valor U correspondiente a la segunda muestra es 2.

Como $N_1 = 2$ y $N_2 = 3$, estos valores satisfacen $U_1 + U_2 = N_1N_2$; es decir, $4 + 2 = (2)(3) = 6$.

10.10. Una población consta de los valores 7, 12 y 15. De ella se toman dos muestras sin reemplazo: la muestra 1 que consta de un valor y la muestra 2 que consta de dos valores. (Las dos muestras juntas agotan la población.)

a) Determinar la distribución muestral de U.

b) Calcular la media y la varianza de la distribución del inciso *a*).

c) Verificar los resultados del inciso *b*) empleando las fórmulas (3) de este capítulo.

a) Elegimos un muestro sin reemplazo para evitar empates, lo que ocurriría si, por ejemplo, el valor 12 apareciera en las dos muestras.

Existen $3 \cdot 2 = 6$ posibilidades para la selección de las muestras, como se aprecia en la tabla 10-16. Se debe notar que, de igual manera, con facilidad se pueden usar los rangos 1, 2 y 3 en vez de 7, 12 y 15. El valor U de la tabla 10-16 es el que se encuentra para la muestra 1, pero si se usara U para la muestra 2, la distribución sería la misma.

Tabla 10-16

Muestra 1	Muestra 2		U
7	12	15	2
7	15	12	2
12	7	15	1
12	15	7	1
15	7	12	0
15	12	7	0

b) La media y la varianza que se encontraron en la tabla 10-15 están dadas por

$$\mu_U = \frac{2 + 2 + 1 + 1 + 0 + 0}{6} = 1$$

$$\sigma_U^2 = \frac{(2 - 1)^2 + (2 - 1)^2 + (1 - 1)^2 + (1 - 1)^2 + (0 - 1)^2 + (0 - 1)^2}{6} = \frac{2}{3}$$

c) Mediante las fórmulas (3),

$$\mu_U = \frac{N_1N_2}{2} = \frac{(1)(2)}{2} = 1$$

$$\sigma_U^2 = \frac{N_1N_2(N_1 + N_2 + 1)}{12} = \frac{(1)(2)(1 + 2 + 1)}{12} = \frac{2}{3}$$

De acuerdo con la parte *a*).

10.11. *a*) Determinar la distribución muestral de U del problema 10.9 y graficarla.

b) Obtener la media y la varianza de U directamente de los resultados del inciso *a*).

c) Verificar el inciso *b*) con las fórmulas (3) de este capítulo.

a) En este caso hay $5 \cdot 4 \cdot 3 \cdot 2 = 120$ posibilidades para elegir valores de las dos muestras, y el método del problema 10.9 es demasiado laborioso. Para simplificar este procedimiento, es necesario concentrarse en la muestra más pequeña (de tamaño $N_1 = 2$) y las posibles sumas de los rangos, R_1. La suma de los rangos de la muestra 1 es la *más pequeña* cuando la muestra contiene los dos números con los rangos más bajos (1,2); entonces, $R_1 = 1 + 2 = 3$. Asimismo, la suma de los rangos de la muestra 1 es *la más grande* cuando la muestra consta de los dos números de los rangos más altos (4,5); entonces, $R_1 = 4 + 5 = 9$. Por tanto, R_1 varía de 3 a 9.

La columna 1 de la tabla 10-17 presenta estos valores de R_1 (de 3 a 9) y la columna 2 indica los valores de la muestra 1 correspondientes, cuya suma es R_1. La columna 3 proporciona la frecuencia (o número) de las muestras con la suma R_1; por ejemplo, existen $f = 2$ muestras con $R_1 = 5$. Puesto que $N_1 = 2$ y $N_2 = 3$, tenemos

$$U = N_1N_2 + \frac{N_1(N_1 + 1)}{2} - R_1 = (2)(3) + \frac{(2)(3)}{2} - R_1 = 9 - R_1$$

La probabilidad de que $U = R_1$ (esto es, $\Pr\{U = R_1\}$) se muestra en la columna 5 de la tabla 10-17 y se obtiene al encontrar la frecuencia relativa. La frecuencia relativa se encuentra dividiendo cada frecuencia f entre la suma de todas las frecuencias, o 10. Por ejemplo, $\Pr\{U = 5\} = \frac{2}{10} = 0.2$.

Tabla 10-17

R_1	Valores de la muestra 1	f	U	$\Pr\{U = R_1\}$
3	(1, 2)	1	6	0.1
4	(1, 3)	1	5	0.1
5	(1, 4), (2, 3)	2	4	0.2
6	(1, 5), (2, 4)	2	3	0.2
7	(2, 5), (3, 4)	2	2	0.2
8	(3, 5)	1	1	0.1
9	(4, 5)	1	0	0.1

b) A partir de las columnas 3 y 4 de la tabla 10-17, tenemos

$$\mu_U = \bar{U} = \frac{\sum fU}{\sum f} = \frac{(1)(6) + (1)(5) + (2)(4) + (2)(3) + (2)(2) + (1)(1) + (1)(0)}{1 + 1 + 2 + 2 + 2 + 1 + 1} = 3$$

$$\sigma_U^2 = \frac{\sum f(U - \bar{U})^2}{\sum f}$$

$$= \frac{(1)(6-3)^2 + (1)(5-3)^2 + (2)(4-3)^2 + (2)(3-3)^2 + (2)(2-3)^2 + (1)(1-3)^2 + (1)(0-3)^2}{10}$$

$$= 3$$

Otro método

$$\sigma_U^2 = \overline{U^2} - \bar{U}^2 = \frac{(1)(6)^2 + (1)(5)^2 + (2)(4)^2 + (2)(3)^2 + (2)(2)^2 + (1)(1)^2 + (1)(0)^2}{10} - (3)^2 = 3$$

c) Mediante las fórmulas (3), con $N_1 = 2$ y $N_2 = 3$, tenemos

$$\mu_U = \frac{N_1N_2}{2} = \frac{(2)(3)}{2} = 3 \qquad \sigma_U^2 = \frac{N_1N_2(N_1 + N_2 + 1)}{12} = \frac{(2)(3)(6)}{12} = 3$$

10.12. Si N números en un conjunto se clasifican de 1 a N, demostrar que la suma de los rangos es $[N(N + 1)]/2$.

Sea R la suma de los rangos. Entonces, tenemos

$$R = 1 + 2 + 3 + \cdots + (N - 1) + N \tag{16}$$

$$R = N + (N - 1) + (N - 2) + \cdots + 2 + 1 \tag{17}$$

donde la suma en la ecuación (17) se obtiene al escribir la suma de (16) al revés. Cuando se suman las ecuaciones (16) y (17), resulta

$$2R = (N + 1) + (N + 1) + (N + 1) + \cdots + (N + 1) + (N + 1) = N(N + 1)$$

puesto que $(N + 1)$ ocurre N veces en la suma; por tanto, $R = [N(N + 1)]/2$. Es posible llegar a esta expresión utilizando el resultado del álgebra básica de progresiones y series aritméticas.

10.13. Si R_1 y R_2 son las sumas respectivas de los rangos para las muestras 1 y 2 en la prueba U, demostrar que $R_1 + R_2 = [N(N + 1)]/2$.

Se supone que no existen empates en los datos muestrales. Entonces R_1 debe ser la suma de algunos de los rangos (números) del conjunto 1, 2, 3, . . . , N, mientras que R_2 tiene que ser la suma de los rangos restantes del conjunto. Así, la suma $R_1 + R_2$ tiene que ser la suma de todos los rangos del conjunto; esto es, $R_1 + R_2 = 1 + 2 + 3 \cdots + N = [N(N + 1)]/2$ según el problema 10.12.

PRUEBA *H* DE KRUSKAL-WALLIS

10.14. Una compañía desea comprar una de cinco máquinas: *A, B, C*, *D* o *E*. En un experimento diseñado para determinar si hay diferencia en el desempeño entre las máquinas, cinco operadores experimentados trabajan en ellas durantes tiempos iguales. La tabla 10-18 muestra el número de las unidades que produce cada máquina. Probar la hipótesis de que no existe diferencia entre las máquinas a un nivel de significancia de *a*) 0.05, *b*) 0.01.

Tabla 10-18

A	68	72	77	42	53
B	72	53	63	53	48
C	60	82	64	75	72
D	48	61	57	64	50
E	64	65	70	68	53

Tabla 10-19

						Suma de filas
A	17.5	21	24	1	6.5	70
B	21	6.5	12	6.5	2.5	48.5
C	10	25	14	23	21	93
D	2.5	11	9	14	4	40.5
E	14	16	19	17.5	6.5	73

Puesto que hay cinco muestras (*A, B, C, D* y *E*), $k = 5$. Además, debido a que cada muestra contiene cinco valores, se tienen $N_1 = N_2 = N_3 = N_4 = N_5 = 5$ y $N = N_1 + N_2 + N_3 + N_4 + N_5 = 25$. Arreglando todos los valores en orden de magnitud creciente y asignando rangos adecuados a los empates se sustituye la tabla 10-18 por la 10.19, cuya columna del lado derecho muestra la suma de los rangos. De la tabla 10-19 se deduce que $R_1 = 70$, $R_2 = 48.5$, $R_3 = 93$, $R_4 = 40.5$ y $R_5 = 73$. Así,

$$H = \frac{12}{N(N + 1)} \sum_{j=1}^{k} \frac{R_j^2}{N_j} - 3(N + 1)$$

$$= \frac{12}{(25)(26)} \left[\frac{(70)^2}{5} + \frac{(48.5)^2}{5} + \frac{(93)^2}{5} + \frac{(40.5)^2}{5} + \frac{(73)^2}{5} \right] - 3(26) = 6.44$$

Para $k - 1 = 4$ grados de libertad a un nivel de significancia de 0.05, de acuerdo con el apéndice E, se tiene $\chi^2_{0.95} = 9.49$. Puesto que $6.44 < 9.49$, no se puede rechazar la hipótesis de que no existe diferencia entre las máquinas a un nivel de 0.05, por lo que es claro que no puede rechazarse a un nivel de 0.01. En otras palabras es posible aceptar la hipótesis (o reservarse la opinión) de que no hay diferencia entre las máquinas a los dos niveles.

Observemos que ya se ha resuelto este problema al usar el análisis de varianza (vea el problema 9.8) y se ha llegado a la misma conclusión.

10.15. Resolver el problema 10.14 si se hace la corrección de los empates.

La tabla 10-20 muestra el número de empates que corresponde a cada una de las observaciones empatadas. Por ejemplo, 48 ocurre dos veces, donde $T = 2$, y 53 ocurre cuatro veces, donde $T = 4$. Al calcular $T^3 - T$ para cada uno de estos valores de T y sumando, se encuentra que $\sum(T^3 - T) = 6 + 60 + 24 + 6 + 24$, como se muestra en la tabla 10-20. Entonces, puesto que $N = 25$, el factor de corrección es

$$1 - \frac{\sum(T^3 - T)}{N^3 - N} = 1 - \frac{120}{(25)^3 - 25} = 0.9923$$

Tabla 10-20

Observación	48	53	64	68	72	
Número de empates (T)	2	4	3	2	3	
$T^3 - T$	6	60	24	6	24	$\sum(T^3 - T) = 120$

y el valor corregido de H es

$$H_c = \frac{6.44}{0.9923} = 6.49$$

Esta corrección no es suficiente para cambiar la decisión que se tomó en el problema 10.14.

10.16. Se eligen tres muestras de manera aleatoria de una población. Arreglando los datos de acuerdo con su rango, resulta la tabla 10-21. Determinar si existe alguna diferencia entre las muestras a un nivel de significancia de *a*) 0.05, *b*) 0.01.

Tabla 10-21

Muestra 1	7	4	6	10	
Muestra 2	11	9	12		
Muestra 3	5	1	3	8	2

Aquí $k = 3$, $N_1 = 4$, $N_2 = 3$, $N_3 = 5$, $N = N_1 + N_2 + N_3 = 12$, $R_1 = 7 + 4 + 6 + 10 = 27$, $R_2 = 11 + 9 + 12 = 32$ y $R_3 = 5 + 1 + 3 + 8 + 2 = 19$. Por tanto,

$$H = \frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{N_j} - 3(N+1) = \frac{12}{(12)(13)}\left[\frac{(27)^2}{4} + \frac{(32)^2}{3} + \frac{(19)^2}{5}\right] - 3(13) = 6.83$$

a) Con $k - 1 = 3 - 1 = 2$ grados de libertad, $\chi^2_{0.95} = 5.99$. En consecuencia, puesto que $6.83 > 5.99$, es posible concluir que existe una diferencia entre las muestras a un nivel de 0.05.

b) Para 2 grados de libertad, $\chi^2_{0.95} = 9.21$. En consecuencia, puesto que $6.83 < 9.21$, no se puede concluir que hay una diferencia entre las muestras a un nivel de 0.01.

PRUEBA DE CORRIDAS DE ALEATORIEDAD

10.17. En 30 lanzamientos de una moneda se obtiene la siguiente secuencia de caras (H) y cruces (T):

H T T H T H H H T H H T T H T

H T H H T H T T H T H H T H T

a) Determinar el número de corridas, V.

b) Probar a un nivel de significancia de 0.05 si la secuencia es aleatoria.

a) Utilizando una barra vertical para indicar una corrida, se ve de

H | T T | H | T | H H H | T | H H | T T | H | T |

H | T | H H | T | H | T T | H | T | H H | T | H | T |

que el número de corridas es $V = 22$.

b) Existen $N_1 = 16$ caras y $N_2 = 14$ cruces en la muestra de lanzamientos dada, y a partir del inciso *a*), el número de corridas es $V = 22$. Por tanto, de las fórmulas (13) de este capítulo, se tiene

$$\mu_V = \frac{2(16)(14)}{16 + 14} + 1 = 15.93 \qquad \sigma_V^2 = \frac{2(16)(14)[2(16)(14) - 16 - 14]}{(16 + 14)^2(16 + 14 - 1)} = 7.175$$

o $\sigma_V = 2.679$. Así, el puntaje z que corresponde a $V = 22$ corridas es

$$Z = \frac{V - \mu_V}{\sigma_V} = \frac{22 - 15.93}{2.679} = 2.27$$

Ahora, en el caso de una prueba de dos colas a un nivel de significancia de 0.05, se debería aceptar la hipótesis H_0 del carácter aleatorio si $-1.96 \leq z \leq 1.96$ y se debería rechazar en caso contrario (vea la figura 10-4). Puesto que el valor calculado de z es $2.27 > 1.96$ se concluye que los lanzamientos no son aleatorios a un nivel de 0.05. Vemos que existen *demasiadas* corridas, lo que indica un *patrón cíclico* en los lanzamientos.

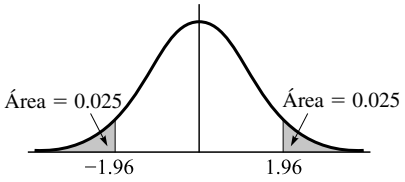


Figura 10-4

Si se usa una corrección de continuidad, el puntaje z anterior se sustituye por

$$z = \frac{(22 - 0.5) - 15.93}{2.679} = 2.08$$

y se llega a la misma conclusión.

10.18. Una máquina produce una muestra de 48 herramientas, la cual se indica en la siguiente secuencia de herramientas buenas (B) y defectuosas (D).

B B B B B B D D B B B B B B B B

B B D D D D B B B B B B D B B B

B B B B B B D D B B B B B D B B

Probar el carácter aleatorio de la secuencia a un nivel de significancia de 0.05.

Los números de D y B son $N_1 = 10$ y $N_2 = 38$, respectivamente, y el número de corridas es $V = 11$. Por tanto, la media y la varianza están dadas por

$$\mu_V = \frac{2(10)(38)}{10 + 38} + 1 = 16.83 \qquad \sigma_V^2 = \frac{2(10)(38)[2(10)(38) - 10 - 38]}{(10 + 38)^2(10 + 38 - 1)} = 4.997$$

por lo que $\sigma_V = 2.235$.

Para una prueba de dos colas a un nivel de 0.05 se debería aceptar la hipótesis H_0 del carácter aleatorio si $-1.96 \leq z \leq 1.96$ (vea la figura 10-4), en caso contrario se rechaza. Como el puntaje z que corresponde a $V = 11$ es

$$Z = \frac{V - \mu_V}{\sigma_V} = \frac{11 - 16.83}{2.235} = -2.61$$

y $-2.61 < -1.96$, es posible rechazar H_0 a un nivel de 0.05.

La prueba muestra que existen *poquísimas* corridas, lo que indica un agrupamiento de herramientas defectuosas. En otras palabras, parece haber un *patrón de tendencia* en la producción de herramientas defectuosas, por lo que es recomendable una revisión adicional del proceso de producción.

10.19. *a*) Generar todas las secuencias posibles que constan de tres *a* y dos *b,* y proporcionar los números de corridas, *V*, que corresponden a cada secuencia.
b) Determinar la distribución muestral de *V*.
c) Calcular la distribución de probabilidad de *V*.

a) El número de posibles secuencias que constan de tres *a* y dos *b* es

$$\binom{5}{2} = \frac{5!}{2!3!} = 10$$

Estas secuencias se muestran en la tabla 10-22, junto con el número de corridas correspondiente a cada secuencia.

Tabla 10-22

Secuencia					Corridas (V)
a	*a*	*a*	*b*	*b*	2
a	*a*	*b*	*a*	*b*	4
a	*a*	*b*	*b*	*a*	3
a	*b*	*a*	*b*	*a*	5
a	*b*	*b*	*a*	*a*	3
a	*b*	*a*	*a*	*b*	4
b	*b*	*a*	*a*	*a*	2
b	*a*	*b*	*a*	*a*	4
b	*a*	*a*	*a*	*b*	3
b	*a*	*a*	*b*	*a*	4

Tabla 10-23

V	f
2	2
3	3
4	4
5	1

b) La distribución muestral de *V* está dada en la tabla 10-23 (que se obtuvo de la tabla 10-21), donde *V* denota el número de corridas y *f* la frecuencia. Por ejemplo, la tabla 10-23 muestra que hay un 5, cuatro 4, etcétera.
c) La distribución de probabilidad de *V* se obtiene de la tabla 10-23 al dividir cada frecuencia por la frecuencia total $2 + 3 + 4 + 1 = 10$. Por ejemplo, $\Pr\{V = 5\} = \frac{1}{10} = 0.1$.

10.20. Calcular *a*) la media, *b*) la varianza del número de corridas del problema 10.19 a partir de los resultados obtenidos.

a) A partir de la tabla 10-22 se tiene

$$\mu_V = \frac{2 + 4 + 3 + 5 + 3 + 4 + 2 + 4 + 3 + 4}{10} = \frac{17}{5}$$

Otro método
De la tabla 10-22, con el método de datos agrupados, obtenemos

$$\mu_V = \frac{\sum fV}{\sum f} = \frac{(2)(2) + (3)(3) + (4)(4) + (1)(5)}{2 + 3 + 4 + 1} = \frac{17}{5}$$

b) Empleando el método de datos agrupados para calcular la varianza, a partir de la tabla 10-23 obtenemos

$$\sigma_V^2 = \frac{\sum f(V - \bar{V})^2}{\sum f} = \frac{1}{10}\left[(2)\left(2 - \frac{17}{5}\right)^2 + (3)\left(3 - \frac{17}{5}\right)^2 + (4)\left(4 - \frac{17}{5}\right)^2 + (1)\left(5 - \frac{17}{2}\right)^2\right] = \frac{21}{25}$$

Otro método

Como en el capítulo 5, la varianza está dada por

$$\sigma_V^2 = \bar{V}^2 - \bar{V}^2 = \frac{(2)(2)^2 + (3)(3)^2 + (4)(4)^2 + (1)(5)^2}{10} - \left(\frac{17}{5}\right)^2 = \frac{21}{25}$$

10.21. Resolver el problema 10.21 con base en las fórmulas (13) de este capítulo.

Puesto que hay tres a y dos b, se tiene $N_1 = 3$ y $N_2 = 2$. Por tanto,

a) $$\mu_V = \frac{2N_1N_2}{N_1 + N_2} + 1 = \frac{2(3)(2)}{3 + 2} + 1 = \frac{17}{5}$$

b) $$\sigma_V^2 = \frac{2N_1N_2(2N_1N_2 - N_1 - N_2)}{(N_1 + N_2)^2(N_1 + N_2 - 1)} = \frac{2(3)(2)[2(3)(2) - 3 - 2]}{(3 + 2)^2(3 + 2 - 1)} = \frac{21}{25}$$

APLICACIONES ADICIONALES DE LA PRUEBA DE CORRIDAS

10.22. Con referencia al problema 10.3, y suponiendo un nivel de significancia de 0.05, determinar si la muestra de tiempos de vida de las pilas que produce la compañía PQR es aleatoria.

La tabla 10-24 muestra los tiempos de vida de las pilas en orden de magnitud creciente. Puesto que en la tabla hay 24 entradas, la mediana se obtiene de las dos entradas centrales, 253 y 262, como $\frac{1}{2}(253 + 262) = 257.5$. Reescribiendo los datos de la tabla 10-23 y utilizando el símbolo a si la entrada está arriba de la mediana y el b si está abajo, se obtiene la tabla 10-25, en la cual hay 12 a, 12 b y 15 corridas. Por tanto, $N_1 = 12$, $N_2 = 12$, $N = 24$, $V = 15$, y se tiene

$$\mu_V = \frac{2N_1N_2}{N_1 + N_2} + 1 = \frac{2(12)(12)}{12 + 12} + 1 = 13 \qquad \sigma_V^2 = \frac{2(12)(12)(264)}{(24)^2(23)} = 5.739$$

por lo que $$Z = \frac{V - \mu_V}{\sigma_V} = \frac{15 - 13}{2.396} = 0.835$$

Empleando una prueba de dos colas a un nivel de significancia de 0.05, se debería aceptar la hipótesis de aleatoriedad si $-1.96 \leq z \leq 1.96$. Puesto que 0.835 está dentro de este rango, se concluye que la muestra es aleatoria.

Tabla 10-24

198	211	216	219	224	225	230	236
243	252	253	253	262	264	268	271
272	275	282	284	288	291	294	295

Tabla 10-25

a	*b*	*b*	*a*	*a*	*b*	*a*	*b*
b	*b*	*a*	*a*	*b*	*a*	*b*	*b*
a	*a*	*b*	*b*	*a*	*b*	*a*	*a*

10.23. Resolver el problema 10.5 usando la prueba de corridas del carácter aleatorio.

El arreglo de todos los valores de las dos muestras aparece en la línea 1 de la tabla 10-8. Usando los símbolos a y b para los datos de las muestras I y II, respectivamente, el arreglo se convierte en

$$b\ b\ b\ b\ b\ b\ b\ b\ a\ a\ a\ a\ a\ b\ b\ a\ a\ a$$

Puesto que hay cuatro corridas, se tiene $V = 4$, $N_1 = 8$ y $N_2 = 10$. Entonces,

$$\mu_V = \frac{2N_1N_2}{N_1 + N_2} + 1 = \frac{2(8)(10)}{18} + 1 = 9.889$$

$$\sigma_V^2 = \frac{2N_1N_2(2N_1N_2 - N_1 - N_2)}{(N_1 + N_2)^2(N_1 + N_2 - 1)} = \frac{2(8)(10)(142)}{(18)^2(17)} = 4.125$$

de modo que $$Z = \frac{V - \mu_V}{\sigma_V} = \frac{4 - 9.889}{2.031} = -2.90$$

Si H_0 es la hipótesis de que no existe una diferencia entre las aleaciones, es también la hipótesis de que la secuencia anterior es aleatoria. Se aceptaría esta hipótesis si $-1.96 \leq z \leq 1.96$, y en caso contrario se debería rechazar. Puesto que $Z = -2.90$ se ubica fuera de este intervalo, se rechaza H_0 y se llega a la misma conclusión que la del problema 10.5.

Observemos que si se hace una corrección de continuidad,

$$Z = \frac{V - \mu_V}{\sigma_V} = \frac{(4 + 0.5) - 9.889}{2.031} = -2.65$$

se llega a la misma conclusión.

CORRELACIÓN DE RANGOS

10.24. La tabla 10-6 muestra cómo 10 estudiantes, ordenados alfabéticamente, se clasificaron de acuerdo con sus logros académicos tanto en el laboratorio como en la clase de un curso de biología. Determinar el coeficiente de correlación de rangos.

Tabla 10-26

Laboratorio	8	3	9	2	7	10	4	6	1	5
Lectura	9	5	10	1	8	7	3	4	2	6

La diferencia entre los rangos, D, en el laboratorio y en la clase de cada estudiante está dada en la tabla 10-27, la cual también proporciona D^2 y $\sum D^2$. Así,

$$r_s = 1 - \frac{6\sum D^2}{N(N^2 - 1)} = 1 - \frac{6(24)}{10(10^2 - 1)} = 0.8545$$

lo que indica que hay una relación notable entre los logros en las secciones de laboratorio y clase del curso.

Tabla 10-27

Diferencia de filas (D)	-1	-2	-1	1	-1	3	1	2	-1	-1	
D^2	1	4	1	1	1	9	1	4	1	1	$\sum D^2 = 24$

10.25. La tabla 10-28 muestra las estaturas de una muestra de 12 padres y sus hijos adultos mayores. Calcular el coeficiente de la correlación de rangos.

Tabla 10-28

Altura del padre (pulgadas)	65	63	67	64	68	62	70	66	68	67	69	71
Altura del hijo (pulgadas)	68	66	68	65	69	66	68	65	71	67	68	70

Ordenadas en magnitud creciente, las estaturas de los padres son

$$62 \quad 63 \quad 64 \quad 65 \quad 66 \quad 67 \quad 68 \quad 69 \quad 70 \quad 71 \tag{18}$$

Puesto que los lugares sexto y séptimo de este arreglo representan la misma estatura (67 pulgadas), se asigna un rango medio $\frac{1}{2}(6 + 7) = 6.5$ a estos lugares. De manera similar, a los lugares octavo y noveno se les asigna el rango $\frac{1}{2}(8 + 9) = 8.5$. Por tanto, a las estaturas de los padres se asignan los rangos

$$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6.5 \quad 6.5 \quad 8.5 \quad 8.5 \quad 10 \quad 11 \quad 12 \tag{19}$$

De igual manera, arregladas en el orden de magnitud ascendente, las estaturas de los hijos son

$$65 \quad 65 \quad 66 \quad 66 \quad 67 \quad 68 \quad 68 \quad 68 \quad 68 \quad 69 \quad 70 \quad 71 \tag{20}$$

y puesto que los lugares sexto, séptimo, octavo y noveno representan la misma estatura (68 pulgadas), se les asigna el rango medio $\frac{1}{4}(6 + 7 + 8 + 9) = 7.5$. Por tanto, a las estaturas de los hijos se les asignan los rangos

$$1.5 \quad 1.5 \quad 3.5 \quad 3.5 \quad 5 \quad 7.5 \quad 7.5 \quad 7.5 \quad 7.5 \quad 10 \quad 11 \quad 12 \tag{21}$$

Usando las correspondencias entre (18) y (19), y (20) y (21), se puede sustituir la tabla 10-28 por la 10-29. La tabla 10-30 muestra la diferencia entre los rangos, D, y los cálculos de D^2 y $\sum D^2$, donde

$$r_s = 1 - \frac{6\sum D^2}{N(N^2 - 1)} = 1 - \frac{6(72.50)}{12(12^2 - 1)} = 0.7465$$

El resultado concuerda bien con el coeficiente de correlación obtenido con otros métodos (vea los problemas 8.26, 8.28, 8.30 y 8.32).

Tabla 10-29

Fila de padre	4	2	6.5	3	8.5	1	11	5	8.5	6.5	10	12
Fila de hijo	7.5	3.5	7.5	1.5	10	3.5	7.5	1.5	12	5	7.5	11

Tabla 10-30

D	−3.5	−1.5	−1.0	1.5	−1.5	−2.5	3.5	3.5	−3.5	1.5	2.5	1.0
D^2	12.25	2.25	1.00	2.25	2.25	6.25	12.25	12.25	12.25	2.25	6.25	1.00

$\sum D^2 = 72.50$

PROBLEMAS SUPLEMENTARIOS

LA PRUEBA DE LOS SIGNOS

10.26. Una compañía afirma que si se agrega su producto al tanque de gasolina de un automóvil, las millas por galón aumentarán. Para probar esta afirmación, se seleccionan 15 automóviles diferentes y se miden las millas por galón con y sin el aditivo. Los resultados se muestran en la tabla 10-31. Suponiendo que las condiciones de manejo son iguales determine si existe una diferencia debida al aditivo para niveles de significancia de *a*) 0.05, *b*) 0.01.

Tabla 10-31

Con aditivo	34.7	28.3	19.6	25.1	15.7	24.5	28.7	23.5	27.7	32.1	29.6	22.4	25.7	28.1	24.3
Sin aditivo	31.4	27.2	20.4	24.6	14.9	22.3	26.8	24.1	26.2	31.4	28.8	23.1	24.0	27.3	22.9

10.27. ¿Es posible concluir a un nivel de significancia de 0.05 que las millas por galón del problema 10.26 son *mayores* con el aditivo que sin él?

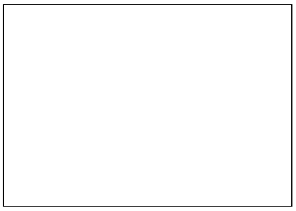
10.28. Un club anuncia un programa para bajar de peso que lo reducirá en por lo menos 6% en un mes si se sigue con rigidez. Para probar la afirmación del club, 36 adultos participan en el programa. De ellos, 25 logran el peso deseado, 6 suben de peso y los demás lo mantienen casi igual. A un nivel de significancia de 0.05, determine si el programa es eficaz.

10.29. Un gerente de capacitación afirma que dando un curso especial al personal de ventas, éstas aumentarán anualmente. Para probar esta afirmación, se da el curso a 24 personas, de las cuales 16 aumentan las ventas, 6 las reducen y las 2 restantes se mantienen sin cambio. Pruebe a un nivel de significancia de 0.05 la hipótesis de que el curso aumentó las ventas de la compañía.

10.30. La compañía refresquera MW establece una "prueba de sabor" en 27 localidades del país para determinar la preferencia relativa del público por dos marcas de refresco, *A* y *B*. En ocho localidades se prefiere la marca *A* sobre la marca *B*. En 17 lugares la marca *B* se prefiere sobre la *A*, y en las localidades restantes no hay una preferencia marcada. ¿Se puede concluir a un nivel de significancia de 0.05 que la marca *B* goza de mayor preferencia que la marca *A*?

10.31. La resistencia a la ruptura de una muestra aleatoria de 25 cuerdas hechas por un fabricante se presenta en la tabla 10-32. Con base en esa muestra, pruebe a un nivel de significancia de 0.05 la afirmación del fabricante de que la resistencia a la ruptura de una cuerda es *a*) 25, *b*) 30, *c*) 35, *d*) 40.

Tabla 10-32



PRUEBA *U* DE MANN-WHITNEY

10.34. Dos profesores —*A* y *B*— enseñan un curso básico de química en la universidad XYZ. En un examen final común, sus estudiantes recibieron las calificaciones que se muestran en la tabla 10-33. Pruebe a un nivel de significancia de 0.05 la hipótesis de que no existe una diferencia entre las calificaciones de los dos profesores.

Tabla 10-33

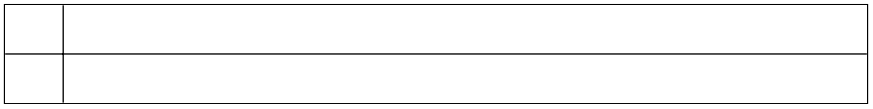


Tabla 10-35

A	30.4	28.7	29.2	32.5	31.7	29.5	30.8	31.1	30.7	31.8
B	33.5	29.8	30.1	31.4	33.8	30.9	31.3	29.6	32.8	33.0

10.39. ¿Se puede usar la prueba *U* para determinar si existe una diferencia entre las máquinas I y II de la tabla 10-1?

10.40. Genere y resuelva un problema con base en la prueba *U*.

10.41. Determine *U* con los datos de la tabla 10-36 usando *a*) el método de fórmulas y *b*) el método de conteo.

Tabla 10-36

Muestra 1	15	25
Muestra 2	20	32

Tabla 10-37

Muestra 1	40	27	30	56
Muestra 2	10	35		

10.42. Resuelva el problema 10.41 con los datos de la tabla 10-37.

10.43. Una población muestra los valores 2, 5, 9 y 12. De esta población se toman dos muestras, de las que la primera contiene uno de estos valores y la segunda los otros tres valores.

a) Determine la distribución muestral de *U* y su gráfica.

b) Calcule la media y la varianza de la distribución anterior, tanto directamente como por medio de las fórmulas.

10.44. Demuestre que $U_1 + U_2 = N_1N_2$.

10.45. Demuestre que $R_1 + R_2 = [N(N + 1)]/2$ en el caso en que el número de empates es *a*) 1, *b*) 2, *c*) cualquier número.

10.46. Si $N_1 = 14$, $N_2 = 12$ y $R_1 = 105$, encuentre *a*) R_2, *b*) U_1, *c*) U_2.

10.47. Si $N_1 = 10$, $N_2 = 16$ y $U_2 = 60$, encuentre *a*) R_1, *b*) R_2, *c*) U_1.

10.48. ¿Cual es el número más grande de los valores N_1, N_2, R_1, R_2, U_1 y U_2 que puede determinarse a partir de los restantes? Justifique su respuesta.

PRUEBA *H* DE KRUSKAL-WALLIS

10.49. Se lleva a cabo un experimento para determinar las producciones de cinco diferentes especies de trigo, *A, B, C, D* y *E*. Cada especie se asigna a cuatro lotes de terreno. Las producciones (en quintales por acre) se muestran en la tabla 10-38. Suponiendo que los lotes tienen fertilidad similar y que las especies se asignan a los lotes de manera aleatoria, determine si existe una diferencia significativa entre las producciones a los niveles de *a*) 0.05, *b*) 0.01.

Tabla 10-38

A	20	12	15	19
B	17	14	12	15
C	23	16	18	14
D	15	17	20	12
E	21	14	17	18

Tabla 10-39

A	33	38	36	40	31	35
B	32	40	42	38	30	34
C	31	37	35	33	34	30
D	27	33	32	29	31	28

10.50. Una compañía desea probar cuatro diferentes tipos de neumáticos: *A, B, C* y *D*. Los tiempos de vida de los neumáticos, determinados por sus bandas (o dibujos), se listan (en miles de millas) en la tabla 10-39. Cada tipo se ha probado en seis automóviles similares a los cuales se asignaron los neumáticos de manera aleatoria. Determine si hay una diferencia significativa entre los neumáticos a los niveles de *a*) 0.05, *b*) 0.01.

10.51. Un maestro desea probar tres diferentes métodos pedagógicos: I, II, III. Para hacerlo, elige de manera aleatoria tres grupos de cinco estudiantes cada uno y enseña a cada grupo de manera diferente. Después, se aplica el mismo examen a todos los estudiantes, y se obtienen las calificaciones que se presentan en la tabla 10-40. Determine a los niveles de significancia de *a*) 0.05 y *b*) 0.01 si existe diferencia entre los métodos pedagógicos.

Tabla 10-40

PRUEBA DE CORRIDAS DE ALEATORIEDAD

10.55. Determine el número de corridas, *V*, de cada una de estas secuencias:

a) A B A B B A A A B B A B

b) H H T H H H T T T T H H T H H T H T

10.56. Se tomó una muestra de 25 individuos para determinar si les gustó o no un producto (sus preferencias están indicadas por S y N, respectivamente). La siguiente secuencia indica la muestra que resulta:

S S N N N N S S S N S N N S N N N N N S S S S N N

a) Determine el número de corridas, *V*.

b) Pruebe a un nivel de significancia de 0.05, si las respuestas son aleatorias.

10.57. Utilice la prueba de corridas en las secuencias (10) y (11) de este capítulo, y enuncie cualquier conclusión sobre el carácter de aleatoriedad.

10.58. *a*) Genere todas las secuencias posibles que consten de dos *a* y una *b*, y proporcione el número de corridas, *V*, que corresponde a cada secuencia.

b) Determine la distribución muestral de *V*.

c) Calcule la distribución de probabilidad de *V*.

10.59. En el problema 10.58 determine la media y la varianza de *V a*) directamente a partir de la distribución muestral y *b*) mediante el método de fórmulas.

10.60. Resuelva los problemas 10.58 y 10.59 en los casos en que existan *a*) dos *a* y dos *b, b*) una *a* y tres *b,* y *c*) una *a* y cuatro *b*.

10.61. Resuelva los problemas 10.58 y 10.59 para los casos en que hay *a*) dos *a* y cuatro *b, b*) tres *a* y tres *b*.

APLICACIONES ADICIONALES DE LA PRUEBA DE CORRIDAS

10.62. Suponga un nivel de significancia de 0.05, para determinar si la muestra de 40 calificaciones en la tabla 10-5 es aleatoria.

10.63. Los precios finales de una acción en 25 días sucesivos se presentan en la tabla 10-42. Determine, a un nivel de significancia de 0.05, si los precios son aleatorios.

Tabla 10-42

10.375	11.125	10.875	10.625	11.500
11.625	11.250	11.375	10.750	11.000
10.875	10.750	11.500	11.250	12.125
11.875	11.375	11.875	11.125	11.750
11.375	12.125	11.750	11.500	12.250

10.64. Los primeros dígitos de $\sqrt{2}$ son 1.41421 35623 73095 0488 $\cdots$. ¿Qué conclusiones se pueden establecer en cuanto al carácter aleatorio de los dígitos?

10.65. ¿A qué conclusiones se puede llegar sobre el carácter aleatorio de los siguientes dígitos?

a) $\sqrt{3} = 1.73205\ 08075\ 68877\ 2935 \cdots$

b) $\pi = 3.14159\ 26535\ 89793\ 2643 \cdots$

10.66. Resuelva el problema 10.30 usando la prueba de corridas del carácter aleatorio.

10.67. Resuelva el problema 10.32 empleando la prueba de corridas del carácter aleatorio.

10.68. Resuelva el problema 10.34 usando la prueba de corridas del carácter aleatorio.

CORRELACIÓN DE RANGOS

10.69. En un concurso se pidió a dos jueces clasificar ocho candidatos (con los números del 1 al 8) en el orden de sus preferencias. Los jueces entregaron las selecciones que se muestran en la tabla 10-43.

a) Estime el coeficiente de la correlación de rangos.

b) Decida en qué medida coinciden los jueces en sus selecciones.

Tabla 10-43

Primer juez	5	2	8	1	4	6	3	7
Segundo juez	4	5	7	3	2	8	1	6

10.70. El coeficiente de correlación de rangos se deduce usando los datos clasificados de la fórmula momento-producto del capítulo 8. Ilustre esta afirmación usando los dos métodos para resolver un problema.

10.71. ¿Es posible determinar el coeficiente de correlación de rangos de datos agrupados? Explique la respuesta e ilústrela con un ejemplo.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

10.26. Existe una diferencia a un nivel de 0.05, pero no a uno de 0.01. **10.27.** Sí.

10.28. El programa es eficaz a un nivel de 0.05.

10.29. Es posible rechazar la hipótesis de un aumento de ventas a un nivel de 0.05. **10.30.** No.

10.31. *a*) Rechazar. *b*) Aceptar. *c*) Aceptar. *d*) Rechazar.

10.34. No existe una diferencia a un nivel de 0.05. **10.35.** No.

10.36. *a*) Sí. *b*) Sí. **10.37.** Sí. **10.38.** *a*) Sí. *b*) Sí.

10.41. 3. **10.42.** 6. **10.49.** No hay una diferencia a ningún nivel.

10.50. La diferencia es significativa a un nivel de 0.05 pero no a uno de 0.01.

10.51. La diferencia es significativa a un nivel de 0.05 pero no a uno de 0.01.

10.52. Hay una diferencia significativa entre las calificaciones a los dos niveles.

10.55. *a*) 8. *b*) 10. **10.56.** *a*) 10. *b*) Las respuestas son aleatorias a un nivel de 0.05.

10.62. La muestra no es aleatoria a un nivel de 0.05. Existen *demasiadas* corridas, lo que indica un patrón cíclico.

10.63. La muestra no es aleatoria a un nivel de 0.05. Existen muy pocas corridas, lo que indica un patrón de tendencia.

10.64. Los dígitos son aleatorios a un nivel de 0.05.

10.65. *a*) Los dígitos son aleatorios a un nivel de 0.05. *b*) Los dígitos son aleatorios a un nivel de 0.05.

10.69. *a*) 0.67. *b*) Los jueces no coincidieron de manera significativa en sus selecciones.

PROBLEMAS APORTADOS

1. Una empresa de investigación de mercado envía a un nuevo empleado para levantar encuestas de preferencia de dos distintas marcas de refrescos de cola (a y b). Las encuestas que entregó el empleado se resumen en la siguiente tabla. Se tienen indicios previos de que en la realidad no hay una marcada preferencia entre el público por alguna de las marcas.

Núm. de cliente	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Marca preferida	a	b	a	a	b	b	a	b	a	b	a	b	a	a
Signos	+	−	+	+	−	−	+	−	+	−	+	−	+	+

Usar la prueba de los signos al nivel de significancia de 0.05 para verificar si efectivamente se puede considerar que no hay diferencia en la preferencia por la marca a o la marca b.

2. En el ejercicio anterior, el gerente de la empresa sospecha que el empleado ha falseado las encuestas. Use la prueba de corridas al nivel de significancia de 0.05 para verificar si puede considerarse aleatoriedad en los datos recabados o, en su caso, se puede sustentar la sospecha del gerente.

Capítulo 11

Métodos bayesianos

Probabilidad subjetiva

Los métodos estadísticos que se desarrollaron hasta ahora se basan totalmente en los enfoques clásicos y de frecuencia de la probabilidad (vea la página 5). Por otro lado, los métodos bayesianos dependen también de una tercera perspectiva de probabilidad, que se conoce como subjetiva o personal.

Resulta central para los métodos bayesianos el proceso de asignar probabilidades a los parámetros, las hipótesis y los modelos, así como actualizar estas probabilidades con base en los datos observados. Por ejemplo, los bayesianos no tratan θ, la media de una población normal, como una constante desconocida; ellos la consideran como el valor que toma una variable aleatoria, digamos Θ, con una función de densidad de probabilidad sobre la recta real. De manera similar, la hipótesis de que es posible asignar una certeza de 0.3 a la probabilidad de que una moneda no esté cargada refleja el grado de creencia de que la moneda no está cargada.

Por tanto, según el enfoque bayesiano, la propiedad del carácter aleatorio pertenece a las hipótesis, los modelos y las cantidades fijas como los parámetros, y también a cantidades variables y observables como las variables aleatorias convencionales. A menudo, las probabilidades que describen el grado de conocimiento o ignorancia de tales entidades no variables se llaman *probabilidades subjetivas* y, en general, se determinan mediante la intuición y la experiencia previa e independiente de cualesquiera observaciones actuales o futuras. En este libro no se discute el controvertido pero crucial asunto del significado y la medición de las probabilidades subjetivas. En lugar de ello, el enfoque se centra en cómo se utilizan las probabilidades a priori en el tratamiento bayesiano de algunos de los problemas estadísticos que se consideraron antes.

EJEMPLO 11.1 Afirmaciones que implican las probabilidades clásicas son: *a*) la probabilidad de que con un dado no cargado se obtenga un 3 o un 5 son una en tres; *b*) la probabilidad de elegir una ficha roja de una caja que contiene dos fichas rojas y tres verdes es dos en cinco. Ejemplos del enfoque de frecuencia de la probabilidad son: *a*) con base en estadísticas oficiales, es prácticamente cero la probabilidad de que cierta persona de Estados Unidos muera por intoxicación de alimentos en el año próximo; *b*) se lanza una moneda 100 veces y se estima que la probabilidad de que resulte cara es $37/100 = 0.37$. Las afirmaciones que implican probabilidades subjetivas incluyen: *a*) un estudiante está 80% seguro de que recibirá una A en su curso; *b*) se cree que las posibilidades de que haya vida en Marte son 1 en 10; *c*) es igualmente probable que la media de la distribución de Poisson sea 1, 1.5 o 2.

Distribuciones a priori y a posteriori

El siguiente ejemplo es útil para introducir algunos términos comunes en la estadística bayesiana.

EJEMPLO 11.2 Una caja contiene dos monedas no cargadas y una cargada con la probabilidad de que caiga cara igual a $P(H) = 0.2$. Se selecciona al azar una moneda de la caja y se lanza tres veces. Si se obtienen dos caras y una cruz, ¿cuál es la probabilidad del evento F de que la moneda seleccionada no esté cargada, y cuál es la probabilidad del evento B de que la moneda sí lo esté?

Sea D el evento (datos) donde se obtienen dos caras y una cruz en tres lanzamientos. La probabilidad condicional $P(D \mid F)$ de observar los datos con base en la hipótesis de que se lance una moneda no cargada es una probabilidad binomial y se puede obtener a partir de (1) (vea el capítulo 4). De manera similar, es posible obtener la probabilidad condicional

$P(D\,|B)$ de observar D cuando se lanza una moneda cargada. En consecuencia, de acuerdo con el teorema de Bayes (página 8) se tiene

$$P(F\,|D) = \frac{P(D\,|F)P(F)}{P(D\,|F)P(F) + P(D\,|B)P(B)} = \frac{[3(0.5)^3] \cdot \left(\frac{2}{3}\right)}{[3(0.5)^3] \cdot \left(\frac{2}{3}\right) + [3(0.2)^2(0.8)] \cdot \left(\frac{1}{3}\right)} = \frac{250}{282} \approx 0.89$$

También, $P(B\,|D) = 1 - P(F\,|D) \approx 0.11$.

En el contexto bayesiano, en general la probabilidad no condicional $P(F)$ del ejemplo anterior se denomina probabilidad a priori (*antes* de recopilar cualesquiera observaciones) de la hipótesis F de que se lanzó una moneda no cargada, y la probabilidad condicional $P(F\,|D)$ se llama probabilidad a posteriori de la hipótesis F (*después* de observar D). De manera análoga, $P(B)$ y $P(B\,|D)$ son las probabilidades respectivas, a priori y a posteriori, de que se lanzó una moneda cargada. Las probabilidades a priori que se usan aquí son probabilidades clásicas.

El siguiente ejemplo implica una modificación sencilla del ejemplo 11.2 que requiere una extensión del concepto del carácter aleatorio, y pone en juego el concepto de la probabilidad subjetiva.

EJEMPLO 11.3 Una caja contiene un número desconocido de monedas no cargadas y cargadas (con $P(H) = 0.2$ cada una). Se selecciona una moneda de manera aleatoria de la caja y se lanza tres veces. Si se obtienen dos caras y una cruz, ¿cuál es la probabilidad de que la moneda elegida esté cargada?

En el ejemplo 11.2, la probabilidad a priori $P(F)$ de elegir una moneda no cargada se pudo determinar mediante el análisis combinatorio. Puesto que ahora no se conoce la proporción de monedas no cargadas que hay en la caja, no es posible recurrir a $P(F)$ como una probabilidad clásica sin que se acuda a tomas repetidas independientes de la caja y aproximarlas como frecuencia relativa. Por tanto, no es posible aplicar el teorema de Bayes para determinar la probabilidad a posteriori de F.

Sin embargo, los métodos bayesianos proporcionan una solución a la cuestión anterior: proponen, primero, que la probabilidad a priori no conocida $P(F)$ es una cantidad aleatoria, digamos Θ, en virtud de la incertidumbre en cuanto a su valor exacto, y después razonar que es posible llegar a una función de probabilidad o de densidad $\pi(\theta)$ con Θ que refleja el grado de creencia en diversas propuestas acerca de $P(F)$. Por ejemplo, se puede argumentar que en ausencia de cualquier evidencia contradictoria *antes de que se lance la moneda*, es razonable suponer que la caja contiene un número igual de monedas cargadas y no cargadas. Debido a que $P(H) = 0.2$ en el caso de una moneda cargada y 0.5 en el de una no cargada, el parámetro no conocido Θ tendría entonces la función de probabilidad subjetiva a priori que se muestra en la tabla 11-1.

Tabla 11-1

θ	0.2	0.5
$\pi(\theta)$	1/2	1/2

Las distribuciones a priori que proporcionan igual peso a todos los valores posibles de un parámetro son ejemplos a priori *difusos, imprecisos* o *no informativos*, los cuales con frecuencia son recomendables cuando no se dispone de casi ninguna información previa acerca del parámetro. Cuando un parámetro puede tomar cualquier valor en un intervalo finito, la a priori difusa sería, normalmente, la densidad uniforme en ese intervalo. También se encuentran situaciones en las que se usan densidades uniformes a priori sobre toda la recta real; las densidades de este tipo se llaman *impropias*, puesto que el área total bajo ellas es infinita.

A partir de la función de probabilidad a priori de la tabla 11-1, la función de probabilidad a posteriori de Θ después de observar D (dos caras y una cruz en tres lanzamientos), $\pi(\theta\,|D)$, se puede obtener con base en el teorema de Bayes como en el ejemplo 11.2, y se indica en la tabla 11-2 (vea el problema 11.3).

Tabla 11-2

θ	0.2	0.5
$\pi(\theta\,|D)$	32/157	125/157

En este momento es conveniente introducir algunas notaciones que son muy útiles para presentar los métodos bayesianos. Suponga que X es una variable aleatoria con una función de densidad de probabilidad $f(x)$ que depende de

un parámetro no conocido θ. Se supone que la incertidumbre en cuanto al valor de θ se puede representar mediante la función de probabilidad o densidad $\pi(\theta)$ de una variable aleatoria Θ. En consecuencia, es posible considerar la función $f(x)$ como la función de probabilidad, o densidad, condicional de X dada $\Theta = \theta$. Por tanto, se denotará $f(x)$ por $f(x \mid \theta)$ en este capítulo. Además, se denotará la función de probabilidad o densidad conjunta de X y Θ por $f(x; \theta) = f(x \mid \theta)\pi(\theta)$ y la función de probabilidad o densidad a posteriori (o condicional) de Θ dada $X = x$ por $\pi(x \mid \theta)$. Si $x_1, x_2, \ldots, x_n$ es una muestra aleatoria de valores de X, la función de densidad conjunta de la muestra (que también se llama *función de verosimilitud* o *probabilidad*, [vea (19), capítulo 6] se escribirá utilizando la notación de vectores $x = (x_1, x_2, \ldots, x_n)$ como $f(x \mid \theta) = f(x_1 \mid \theta) \cdot f(x_2 \mid \theta) \cdot \cdots f(x_n \mid \theta)$; de manera similar, la función de probabilidad o densidad a posteriori de θ dada la muestra se denotará por $\pi(\theta \mid x)$.

La siguiente versión del teorema de Bayes para variables aleatorias es una consecuencia directa de (26) y (43), del capítulo 2:

$$\pi(\theta \mid \boldsymbol{x}) = \frac{f(\boldsymbol{x}; \theta)}{f(\boldsymbol{x})} = \frac{f(\boldsymbol{x} \mid \theta)\pi(\theta)}{\displaystyle\int_{\Theta} f(\boldsymbol{x} \mid \theta)\pi(\theta)\, d\theta} \tag{1}$$

donde la integral está sobre el rango de valores de Θ y se sustituye por una suma si Θ es discreta.

En las aplicaciones del teorema de Bayes, es poco común que se tenga que efectuar la integración (o la suma) que aparece en el denominador de (1) puesto que su valor es independiente de θ. Por tanto, es posible escribir (1) en la forma

$$\pi(\theta \mid \boldsymbol{x}) \propto f(\boldsymbol{x} \mid \theta)\pi(\theta) \tag{2}$$

que significa que $\pi(\theta \mid x) = C \cdot f(x \mid \theta)\ \pi(\theta)$, donde C es una constante de proporcionalidad que no depende de θ. Una vez que se conoce la forma funcional de la densidad a posteriori se puede determinar la constante "normalizadora" C de modo que $\pi(\theta \mid x)$ se hace una función de densidad de probabilidad. (Vea el ejemplo 11.4.)

Nota 1 Con frecuencia, la convención de usar mayúsculas para designar las variables aleatorias se pasa por alto en las presentaciones bayesianas cuando implican parámetros, práctica que se seguirá de ahora en adelante. Por ejemplo, en el caso que sigue, se usa λ (en vez de Λ) para denotar tanto el parámetro aleatorio como sus posibles valores.

EJEMPLO 11.4 La variable aleatoria X tiene una distribución de Poisson con parámetro desconocido λ. Se ha determinado que λ tiene la función de probabilidad subjetiva a priori dada en la tabla 11-3. Una muestra aleatoria de tamaño 3 produce los valores de X: 2, 0 y 3. Se desea encontrar la distribución a posteriori de λ.

Tabla 11-3

λ	0.5	1.0	1.5
$\pi(\lambda)$	1/2	1/3	1/6

La probabilidad de los datos es $f(\boldsymbol{x} \mid \lambda) = e^{-3\lambda}\dfrac{\lambda^{x_1+x_2+x_3}}{x_1!x_2!x_3!}$. A partir de (1) y (2) se tiene la densidad a posteriori

$$\pi(\lambda \mid \boldsymbol{x}) = \frac{\dfrac{e^{-3\lambda}\lambda^{x_1+x_2+x_3}\pi(\lambda)}{x_1!x_2!x_3!}}{\dfrac{1}{x_1!x_2!x_3!}\displaystyle\sum_{\lambda} e^{-3\lambda}\lambda^{x_1+x_2+x_3}\pi(\lambda)} \propto e^{-3\lambda}\lambda^5\pi(\lambda) \qquad \lambda = 0.5, 1, 1.5$$

La constante de proporcionalidad en la expresión anterior es sólo el recíproco de la suma $\sum_{\lambda} e^{-3\lambda}\lambda^5\pi(\lambda)$ de los tres valores posibles de λ. Cuando se sustituye $\lambda = 0.5, 1.0, 1.5$, respectivamente, y $\pi(\lambda)$ de la tabla 11-3 en la suma anterior, y después se normaliza para que la suma de las probabilidades $\pi(\lambda \mid x)$ sea igual a 1, se obtienen los valores de la tabla 11-4.

Tabla 11-4

λ	0.5	1.0	1.5
$\pi(\lambda \mid \boldsymbol{x})$	0.10	0.49	0.41

EJEMPLO 11.5 La variable aleatoria X tiene una distribución binomial con función de probabilidad dada por

$$f(x|\theta) = \binom{n}{x}\theta^x(1 - \theta)^{n-x} \qquad x = 1, 2, \ldots, n$$

Suponga que no se sabe nada acerca del parámetro θ, por lo que se elige una distribución uniforme (imprecisa) a priori sobre el intervalo [0,1] para θ. Si una muestra de tamaño 4 produce 3 éxitos, se puede obtener la función de densidad de probabilidad a posteriori de θ utilizando (2):

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{\int_\Theta f(\boldsymbol{x}|\theta)\pi(\theta)\,d\theta} = \frac{\binom{4}{3}\theta^3(1 - \theta)\cdot 1}{\int_0^1 \binom{4}{3}\theta^3(1 - \theta)\,d\theta} \propto \theta^3(1 - \theta)$$

La expresión anterior puede reconocerse como una densidad beta [vea (34) del capítulo 4] con $\alpha = 4$ y $\beta = 2$. Puesto que la constante normalizadora aquí debe ser $\frac{1}{B(4, 2)} = \frac{5!}{3!1!}$ (vea el apéndice A), se deduce que la constante de proporcionalidad es 20 y $\pi(\theta|\boldsymbol{x}) = 20\theta^3(1 - \theta)$, $0 < \theta < 1$. Las gráficas de las densidades a priori (uniforme) y a posteriori se muestran en la figura 11-1. La media y la varianza son, respectivamente, 0.5 y $1/12 \approx 0.08$ en el caso de la densidad a priori, mientras que a la densidad a posteriori corresponden $2/3 \approx 0.67$ y $8/252 \approx 0.03$. El corrimiento a la derecha y la concentración aumentada alrededor de la media a medida que se produce un desplazamiento de la densidad a priori a la a posteriori son evidentes en la figura 11-1.

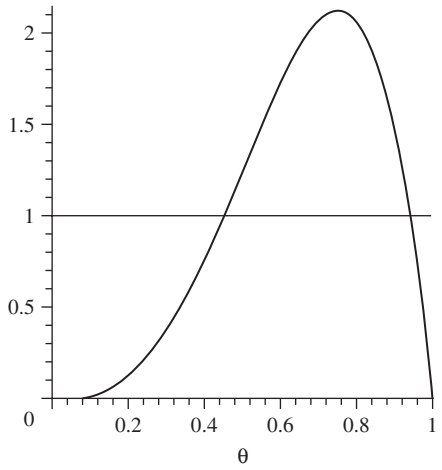


Figura 11-1

MUESTREO A PARTIR DE UNA POBLACIÓN BINOMIAL

El resultado que se obtuvo en el ejemplo 11.5 puede generalizarse de manera sencilla. Suponga que X tiene una distribución binomial con parámetros n y θ [vea (1) del capítulo 4], y que la distribución de probabilidad a priori de θ es beta con una función de densidad [vea (34) del capítulo 4]:

$$\pi(\theta) = \frac{\theta^{\alpha-1}(1 - \theta)^{\beta-1}}{B(\alpha, \beta)} \qquad 0 < \theta < 1 \qquad (\alpha, \beta > 0) \tag{3}$$

donde $B(\alpha, \beta)$ es la función beta (vea el apéndice A). (Observe que si $\alpha = \beta = 1$, entonces $\pi(\theta)$ es la densidad uniforme en [0,1]; esta situación se explica en el ejemplo 11.5.) Entonces, la densidad a posteriori $\pi(\theta|x)$ correspondiente a cualquier valor observado x está dada por

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{\int_\Theta f(\boldsymbol{x}|\theta)\pi(\theta)\,dp} \propto \theta^x(1 - \theta)^{n-x}\cdot\theta^{\alpha-1}(1 - \theta)^{\beta-1} = \frac{\theta^{x+\alpha-1}(1 - \theta)^{n-x+\beta-1}}{B(x + \alpha, n - x + \beta)} \quad 0 < \theta < 1 \tag{4}$$

Ésta puede reconocerse como una densidad beta con parámetros $x + \alpha$ y $n - x + \beta$. Por tanto, se tiene el siguiente teorema:

Teorema 11-1 Si X es una variable aleatoria binomial con parámetros n y θ, y la densidad a priori de θ es beta con parámetros α y β, entonces la densidad a posteriori de θ después de observar $X = x$ es beta con parámetros $x + \alpha$ y $n - x + \beta$.

EJEMPLO 11.6 Suponga que X es binomial con parámetros $n = 10$ y θ desconocida y que $\pi(\theta)$ es beta con parámetros $\alpha = \beta = 2$. Si una observación en X produce $x = 2$, entonces es posible determinar la densidad a posteriori $\pi(\theta|x)$ de la siguiente manera.

A partir del teorema 11-1 se observa que $\pi(\theta|x)$ es beta con parámetros $\alpha = 4$ y $\beta = 10$. Las densidades a priori (simétrica alrededor de 0.5) y a posteriori se muestran en la figura 11-2. Es obvio que el efecto de la observación de la densidad a priori de θ es correr su media de 0.5 hacia abajo a $4/14 \approx 0.29$ y disminuir la varianza de 0.05 a 0.014 [vea (36) del capítulo 4].

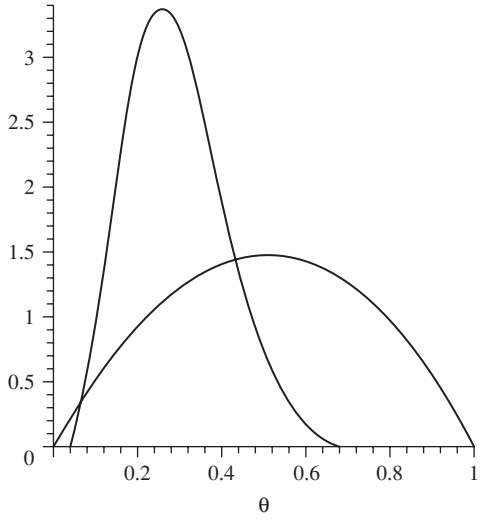


Figura 11-2

MUESTREO A PARTIR DE UNA POBLACIÓN DE POISSON

Teorema 11-2 Si X es una variable aleatoria de Poisson con parámetro λ y la densidad a priori de λ es gamma con parámetros α y β [como en (31), capítulo 4], entonces la densidad a posteriori de λ, dada la muestra $x_1, x_2, \ldots, x_n$, es gamma con parámetros $n\bar{x} + \alpha$ y $\beta/(1 + n\beta)$, donde $\bar{x}$ es la media muestral.

Si $x_1, x_2, \ldots, x_n$ es una muestra de n observaciones en X, entonces la probabilidad de $x = (x_1, x_2, \ldots, x_n)$ se puede escribir como $f(\boldsymbol{x}|\theta) = e^{-n\lambda}\dfrac{\lambda^{n\bar{x}}}{x_1!x_2!\cdots x_n!}$. La densidad a priori de λ está dada por:

$$\pi(\lambda) = \frac{\lambda^{\alpha-1}e^{-\lambda/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \qquad \lambda > 0 \tag{5}$$

Se concluye que la densidad a posteriori de λ es

$$\pi(\lambda|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\lambda)\pi(\lambda)}{\displaystyle\int_{\Lambda} f(\boldsymbol{x}|\lambda)\pi(\lambda)\,d\lambda} \propto \frac{e^{-n\lambda}\lambda^{n\bar{x}} \cdot \lambda^{\alpha-1}e^{-\lambda/\beta}}{\displaystyle\int_0^{\infty} e^{-\lambda}\lambda^{n\bar{x}} \cdot \lambda^{\alpha-1}e^{-\lambda/\beta}d\lambda} = \frac{(1 + n\beta)^{n\bar{x}+\alpha}\lambda^{(n\bar{x}+\alpha)-1}e^{-\lambda(n\beta+1)/\beta}}{\beta^{n\bar{x}+\alpha}\Gamma(n\bar{x} + \alpha)} \quad \lambda > 0 \tag{6}$$

Se puede reconocer la expresión anterior como una densidad gamma, esto prueba el teorema 11-2.

EJEMPLO 11.7 El número de defectos en un carrete de 1000 pies de hilo que fabrica una máquina tiene una distribución de Poisson con media desconocida λ. La distribución a priori de λ es gamma con parámetros $\alpha = 3$ y $\beta = 1$. En total, en una muestra de cinco carretes que se examinaron se encontraron ocho defectos. La distribución a posteriori de λ es gamma con los parámetros $\alpha = 11$ y $\beta = 1/6 \approx 0.17$. Tanto la media como la varianza a priori son ambas 3, mientras que la media y la varianza a posteriori son 1.87 y 0.32, respectivamente. Las dos densidades se muestran en la figura 11-3.

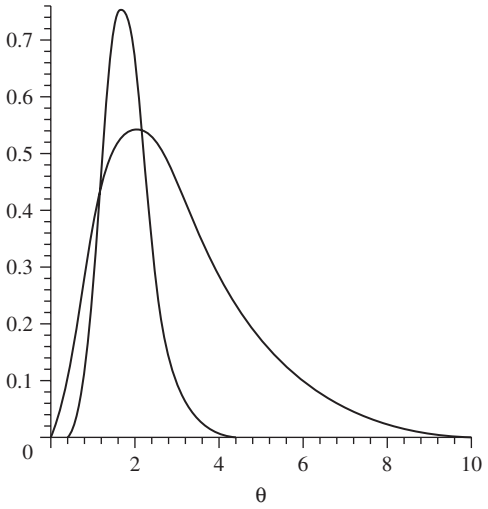


Figura 11-3

MUESTREO A PARTIR DE UNA POBLACIÓN NORMAL CON VARIANZA CONOCIDA

Teorema 11-3 Suponga que una muestra aleatoria de tamaño n se toma de una distribución normal con media θ desconocida y varianza σ^2 conocida. Suponga también que la distribución a priori de θ es normal con media μ y varianza v^2. Entonces, la distribución a posteriori de θ también es normal, con media μ_{post} y varianza v^2_{post} dadas por

$$\mu_{\text{post}} = \frac{\sigma^2\mu + nv^2\bar{x}}{\sigma^2 + nv^2} \qquad v^2_{\text{post}} = \frac{\sigma^2 v^2}{\sigma^2 + nv^2} \tag{7}$$

La probabilidad de las observaciones está dada por

$$f(\boldsymbol{x}\,|\,\theta) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2\right\}$$

Se sabe del problema 5.20 (vea el método 2) que $\sum(x_i - \theta)^2 = \sum(x_i - \bar{x})^2 + n(\bar{x} - \theta)^2$. Si se utiliza la expresión anterior y se hace caso omiso de las constantes multiplicativas que no implican a θ, es posible escribir la probabilidad como

$$f(\boldsymbol{x}\,|\,\theta) \propto \exp\left\{-\frac{n}{2\sigma^2}(\theta - \bar{x})^2\right\}$$

Empleando (2) y el hecho de que $\pi(\theta) = \dfrac{1}{v\sqrt{2\pi}}\exp\left\{-\dfrac{1}{2v^2}(\theta - \mu)^2\right\}$ se tiene la densidad a posteriori de θ como

$$\pi(\theta\,|\,\boldsymbol{x}) \propto \exp\left\{-\frac{1}{2}\left[\frac{n}{\sigma^2}(\theta - \bar{x})^2 + \frac{1}{v^2}(\theta - \mu)^2\right]\right\}$$

Cuando se completa el cuadrado de la expresión entre corchetes, resulta

$$\pi(\theta\,|\,\boldsymbol{x}) \propto \exp\left\{-\frac{[\theta - (\bar{x}v^2 + \mu\sigma^2/n)/(v^2 + \sigma^2/n)]^2}{[2(\sigma^2/n)v^2]/[v^2 + (\sigma^2/n)]}\right\} \qquad -\infty < \theta < \infty$$

Esto prueba que la densidad a posteriori de θ es normal con media y varianza dadas por (7).

Una comparación de las varianzas a priori y a posteriori de θ en el teorema 11-3 destaca algunos hechos importantes. Es conveniente hacer la comparación en términos del recíproco de la varianza, que se conoce como la *precisión* de la distribución o variable aleatoria. Es obvio que cuanto más pequeña sea la varianza de una distribución,

tanto más grande será su precisión. En consecuencia, la precisión es una medida del grado de concentración de una variable aleatoria o qué tan bien se conoce. En el teorema 11-3, si se denota la precisión de las distribuciones a priori y a posteriori de θ, respectivamente, por ξ_{prior} y ξ_{post}, se tiene

$$\xi_{\text{prior}} = \frac{1}{v^2} \qquad \text{y} \qquad \xi_{\text{post}} = \frac{\sigma^2 + nv^2}{\sigma^2 v^2} = \frac{1}{v^2} + \frac{n}{\sigma^2} \tag{8}$$

La cantidad $\frac{n}{\sigma^2}$ puede considerarse como la precisión de los datos (media muestral). Si ésta se denota por ξ_{datos} se tiene el resultado $\xi_{\text{post}} = \xi_{\text{prior}} + \xi_{\text{datos}}$. Esto es, la precisión de la distribución a posteriori es la suma de las precisiones de la a priori y de los datos. También es posible escribir la media a posteriori, dada en (7), en la forma

$$\mu_{\text{post}} = \frac{\sigma^2\mu + nv^2\bar{x}}{\sigma^2 + nv^2} = \frac{\xi_{\text{prior}}\mu + \xi_{\text{datos}}\bar{x}}{\xi_{\text{prior}} + \xi_{\text{datos}}} \tag{9}$$

Esta expresión indica que la media a posteriori es una suma ponderada de la media a priori y de los datos, con los pesos proporcionales a las precisiones respectivas.

Suponga ahora que ξ_{prior} es mucho menor que ξ_{datos}. Entonces, ξ_{post} sería muy cercana a ξ_{datos} y μ_{post} sería muy cercana a $\bar{x}$. En otras palabras, los datos dominarían en ese caso a la información a priori, y la distribución a posteriori resultaría casi proporcional a la probabilidad. En cualquier caso, como se puede ver de (8) y (9), los datos dominarían a la a priori para n muy grande.

EJEMPLO 11.8 Suponga que X está distribuida de manera normal con media θ desconocida y varianza 4 y que $\pi(\theta)$ es normal estándar. Si una muestra de tamaño $n = 10$ produce una media de 0.5, entonces, mediante el teorema 11-3, $\pi(\theta|x)$ es normal con media 0.36 y varianza 0.29. La precisión a posteriori ($\xi_{\text{prior}} = 3.5$) es más de tres veces la precisión a priori ($\xi_{\text{post}} = 1$), lo que es evidente de las densidades que se muestran en la figura 11-4. La precisión de los datos es $10/4 = 2.5$, la cual es razonablemente más grande que la precisión a priori de 1; ello se refleja en que la media a posteriori de 0.36 es más cercana a $\bar{x} = 0.5$ que a la media a priori, 0.

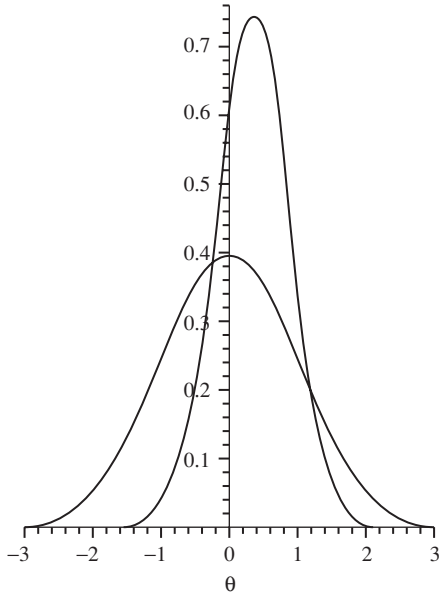


Figura 11-4

DISTRIBUCIONES A PRIORI IMPROPIAS

Las funciones de densidad de probabilidad a priori $\pi(\theta)$ que se han estudiado hasta ahora son todas *propias* en el sentido de que *i)* $\pi(\theta) \geq 0$ y *ii)* $\int_{-\infty}^{\infty} \pi(\theta)\, d\theta = 1$ (vea la página 37). Sin embargo, las densidades a priori que satisfacen la primera condición pero no la segunda debido a que la integral es divergente, se han empleado dentro del marco bayesiano y se conocen como *a priori impropias*. A menudo, surgen como elecciones naturales para representar información a priori imprecisa, o vaga, acerca de los parámetros con rangos infinitos.

Por ejemplo, si se toma la muestra de una población normal con media conocida, digamos 0, pero con varianza θ desconocida, se puede suponer que la densidad a priori de θ está dada por $\pi(\theta) = \frac{1}{\theta}$, $\theta > 0$. Dada una muestra de observaciones $x = (x_1, x_2, \ldots, x_n)$ si se pasa por alto el hecho de que la a priori es impropia y se aplica la fórmula (1), resulta la densidad a posteriori

$$\pi(\theta|\boldsymbol{x}) \propto \frac{1}{\theta^{n/2}} \exp\left\{\frac{-\sum_i x_i^2}{2\theta}\right\} \cdot \frac{1}{\theta} = \theta^{-\frac{n}{2}-1} \exp\left\{\frac{-\sum_i x_i^2}{2\theta}\right\} \qquad \theta > 0 \tag{10}$$

Ésta es una densidad propia que se conoce como una gamma inversa, con parámetros $\alpha = n/2$ y $\beta = \sum_i x_i^2/2$ (vea el problema 11.99). En consecuencia, se ha llegado a una densidad a posteriori propia a partir de una a priori impropia. En realidad, éste será el caso en todas las situaciones con a priori impropias que se encuentran aquí, aunque no siempre es eso lo que ocurre.

EJEMPLO 11.9 Suponga que X es binomial con n conocida y probabilidad de éxito θ desconocida. La densidad a priori para θ dada por $\pi(\theta) = \frac{1}{\theta(1-\theta)}$, $0 < \theta < 1$ es impropia y se llama *a priori de Haldane*. Se pasa por alto el hecho de que $\pi(\theta)$ es impropia y se procede formalmente a obtener la densidad a posteriori $\pi(\theta|x)$ que corresponde a un valor observado x de X:

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int_\Theta f(x|\theta)\pi(\theta)d\theta} \propto \frac{\theta^x(1-\theta)^{n-x}}{\theta(1-\theta)} = \frac{\theta^{x-1}(1-\theta)^{n-x-1}}{B(x, n-x)} \qquad 0 < \theta < 1$$

Se puede observar que la a posteriori es una densidad beta propia con parámetros x y $n - x$.

EJEMPLO 11.10 Suponga que X está distribuida de manera normal con media θ desconocida y varianza σ^2 conocida. En este caso, una distribución a priori impropia para θ está dada por $\pi(\theta) = 1$, $-\infty < \theta < \infty$. Es posible considerar la densidad anterior como la representación de una ignorancia a priori en aquellos intervalos de la misma longitud que tienen el mismo peso sin que importe su ubicación en la línea real. Dado el vector de observación $x = (x_1, x_2, \ldots, x_n)$, la distribución a posteriori de θ bajo esta a priori corresponde a

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)\pi(\theta) \propto \exp\left\{-\frac{\sum_i (x_i - \theta)^2}{2\sigma^2}\right\} \cdot 1 \propto \exp\left\{-\frac{n(\theta - \bar{x})^2}{2\sigma^2}\right\} \qquad -\infty < \theta < \infty$$

que es normal con media $\bar{x}$ y varianza σ^2/n.

DISTRIBUCIONES A PRIORI CONJUGADAS

Observe que los teoremas 11-1, 11-2 y 11-3 comparten una característica importante en términos de que las densidades a priori y a posteriori en cada uno pertenecen a la misma familia de distribuciones. Cuando eso ocurre se dice que la familia de distribuciones a priori que se usa es *conjugada* (o *cerrada*) con respecto a la densidad poblacional $f(x|\theta)$. Por tanto, la familia beta es conjugada con respecto a la distribución binomial (teorema 11-1), la familia gamma es conjugada con respecto a la distribución de Poisson (teorema 11-2) y la familia normal es conjugada con respecto a la distribución normal con varianza conocida (teorema 11-3).

Puesto que $\pi(\theta|\boldsymbol{x}, y) \propto f(y|\theta)\pi(\theta|\boldsymbol{x})$ siempre que $\boldsymbol{x}$ y $\boldsymbol{y}$ sean dos muestras independientes de $f(x|\theta)$, las familias conjugadas hacen más fácil la actualización de densidades a priori de manera secuencial tan sólo cambiando los parámetros de la familia (vea el ejemplo 11.11). En consecuencia, las familias conjugadas son deseables en el análisis bayesiano y existen para la mayoría de las distribuciones que se encuentran comúnmente. No obstante, en la práctica las distribuciones a priori se eligen con base en qué tan bien representan el conocimiento y creencia previos y no por la conveniencia matemática. Sin embargo, si una distribución a priori conjugada se aproxima mucho a una distribución a priori apropiada, pero poco manejable en otro caso, entonces la primera es naturalmente una elección prudente.

A continuación se demuestra que la familia gamma es conjugada con respecto a la distribución exponencial. Suponga que X tiene la densidad exponencial $f(x|\theta) = \theta e^{-\theta x}$, $x > 0$, con θ desconocida, y que la densidad a priori de θ es gamma con parámetros α y β. En ese caso, la densidad a posteriori de θ está dada por

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta) \cdot \pi(\theta) = \frac{\theta^n e^{-\theta n\bar{x}} \theta^{\alpha-1} e^{-\theta/\beta}}{\beta^\alpha \Gamma(\alpha)} = \frac{(1 + n\beta\bar{x})^{n+\alpha} \theta^{n+\alpha-1} e^{-\theta\left(\frac{1}{\beta}+n\bar{x}\right)}}{\beta^{n+\alpha}\Gamma(n+\alpha)} \qquad \theta > 0 \tag{11}$$

Lo anterior establece el siguiente teorema:

Teorema 11-4 Si X tiene la densidad exponencial $f(x|\theta) = \theta e^{-\theta x}$, $x > 0$ con θ desconocida y la densidad a priori de θ es gamma con parámetros α y β, entonces la densidad a posteriori de θ es gamma con parámetros $\alpha + n$ y $\beta/(1 + n\beta\bar{x})$.

EJEMPLO 11.11 En el ejemplo 11.6 suponga que una observación adicional independiente sobre la misma población binomial produce al valor muestral $y = 3$. En consecuencia, se puede estimar la densidad a posteriori $\pi(\theta|x, y)$, ya sea *a*) directamente de la densidad a priori $\pi(\theta)$ dada en el ejemplo 11.6, o *b*) mediante el uso de la densidad a posteriori $\pi(\theta|x)$ que se deduce ahí.

a) Se puede suponer que la densidad a priori es beta con parámetros $\alpha = 2$ y $\beta = 2$, y que un valor muestral de 5 se observa en una variable aleatoria binomial con $n = 20$. En ese caso, el teorema 11-1 da en la densidad a posteriori beta con parámetros $\alpha = 2 + 5 = 7$ y $\beta = 15 + 2 = 17$.

b) Se puede suponer que la densidad a priori es la densidad a posteriori que se obtuvo en el ejemplo 11.6, es decir, una beta con parámetros $\alpha = 4$ y $\beta = 10$, y que se observa el valor muestral de 3 de una variable aleatoria binomial con $n = 10$. El teorema 11-1 proporciona una densidad beta a posteriori con parámetros $\alpha = 4 + 3 = 7$ y $\beta = 7 + 10 = 17$.

EJEMPLO 11.12 Una muestra aleatoria de tamaño n se toma de una distribución geométrica con parámetro θ (vea la página 123): $f(x; \theta) = \theta(1 - \theta)^{x-1}$, $x = 1, 2, \ldots$ Suponga que la densidad a priori de θ es beta con parámetros α y β. En consecuencia, la densidad a posteriori de θ es

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{\int_{\Omega} f(\boldsymbol{x}|\theta)\pi(\theta)dp} \propto \frac{\theta^n(1-\theta)^{n(\bar{x}-1)} \cdot \theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1 \theta^n(1-\theta)^{n(\bar{x}-1)}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta} = \frac{\theta^{\alpha+n-1}(1-\theta)^{\beta+n\bar{x}-n-1}}{B(\alpha+n, \beta+n\bar{x}-n)} \qquad 0 < \theta < 1$$

que también es beta, con parámetros $\alpha + n$ y $\beta + n\bar{x} - n$, donde $\bar{x}$ es la media muestral. En otras palabras, la familia beta es conjugada con respecto a la distribución geométrica.

ESTIMACIÓN PUNTUAL BAYESIANA

Un principio central de los estadísticos bayesianos es que todo lo que se tiene que saber acerca de un parámetro desconocido se encontrará en su distribución a posteriori. Por consiguiente, la estimación puntual bayesiana equivale, básicamente, a encontrar un resumen numérico de la distribución a posteriori del parámetro. A continuación se presentarán algunos resúmenes numéricos que se emplean con este propósito, así como sus méritos relativos en cuanto a qué tan bien representan al parámetro.

EJEMPLO 11.13 En el ejemplo 11.5 se vio que al tomar una muestra de una distribución binomial con una a priori uniforme, la densidad a posteriori de θ es beta con parámetros $\alpha = 4$ y $\beta = 2$; la gráfica de esta densidad se muestra en la figura 11-1. En este caso, un candidato natural para el estatus del resumen de números individuales sería la media de la densidad a posteriori. Se sabe de (36) del capítulo 4 que la media a posteriori está dada por $\alpha/(\alpha + \beta) = 2/3$.

La mediana y la moda (vea la página 83) de la densidad a posteriori son otras dos posibles selecciones como estimaciones puntuales para θ. La moda está dada por $(\alpha - 1)/(\alpha + \beta - 2) = 3/4$ [vea (37) del capítulo 4]. Observe que la moda coincide con el estimador de máxima verosimilitud (vea las páginas 198-199) de θ, a saber, la proporción muestral de éxitos. Como corolario del teorema 11-5 se comprueba que ello es cierto en general en el caso de la distribución binomial con una a priori uniforme.

En este caso la mediana no es atractiva desde una perspectiva práctica puesto que se tiene que determinar numéricamente debido a la falta de una expresión explícita (además de la definición) de la mediana de una distribución beta. Sin embargo, en general, como se verá después, la mediana es en cierto sentido una medición de resúmenes óptima.

El siguiente teorema generaliza algunos de los resultados del ejemplo 11.13.

Teorema 11-5 Si X es una variable aleatoria binomial con parámetros n y θ y la densidad a priori de θ es beta con parámetros α y β, entonces las estimaciones respectivas de θ que proporcionan la media y la moda a posteriori son $\mu_{\text{post}} = (x + \alpha)/(n + \alpha + \beta)$ y $\gamma_{\text{post}} = (x + \alpha - 1)/(n + \alpha + \beta - 2)$.

Comentario 2 Resulta de interés un caso especial del teorema 11-5 cuando tanto α como β equivalen a 1. El estimador de la media a posteriori de θ es, entonces, $(x + 1)/(n + 2)$. En concordancia, si todos los n ensayos resultan ser éxitos (esto es, si $x = n$), entonces la probabilidad de que el siguiente ensayo sea también un éxito está dada por $(n + 1)/(n + 2)$. El resultado anterior tiene una historia respetable y se conoce como *ley de sucesión de Laplace*.

Cuando $\alpha = \beta = 1$ según el teorema 11-5, el estimador de la moda a posteriori γ_{post} de θ se reduce al estimador de máxima verosimilitud x/n. Esta característica también se señaló en el ejemplo 11.13, pero es obvio que el resultado no es válido para α y β en general. Sin embargo, sin que importen los valores de α y β, cuando el tamaño de la muestra es suficientemente grande, tanto μ_{post} como γ_{post} serán cercanas a la proporción muestral x/n. Además, para todo n, μ_{post} es una combinación convexa de la media a priori de θ y de la proporción muestral. (Vea el problema 11.38.)

EJEMPLO 11.14 Suponga que se toma una muestra aleatoria de tamaño n de una distribución normal con media θ desconocida y varianza 1. Suponga también que la distribución a priori de θ es normal, con media 0 y varianza 1. De acuerdo con el teorema 11-3 se advierte que la distribución a posteriori de θ es normal con media $n\bar{x}/(1 + n)$.

Es claro que, en este caso, la media, la mediana y la moda a posteriori son todas ellas idénticas y, por consiguiente, conducen al mismo estimador puntual, $n\bar{x}/(1 + n)$ de θ. En el problema 6.25, página 206, se demostró que el estimador de máxima verosimilitud de θ en este caso es la media muestral $\bar{x}$, que se sabe es insesgada (teorema 5-1). Por otro lado, los estimadores bayesianos obtenidos aquí son sesgados, aunque son insesgados de manera asintótica.

Un resultado general en este sentido, que se deduce con facilidad del teorema 11-6, se presenta a continuación.

Teorema 11-6 Suponga que una muestra aleatoria de tamaño n se toma de una distribución normal con media θ desconocida y varianza σ^2 conocida. También suponga que la distribución a priori de θ es normal con media μ y varianza v^2. Entonces, la media, la mediana y la moda a posteriori proporcionan, todas ellas, el mismo estimador de θ, a saber, $(\sigma^2\mu + nv^2\bar{x})/(\sigma^2 + nv^2)$, donde $\bar{x}$ es la media muestral.

Como se notó en el caso binomial, el estimador de la media a posteriori μ_{post}, que se acaba de obtener, se ubica entre la media a priori μ y el estimador de máxima verosimilitud $\bar{x}$ de θ. Ello puede verse al escribir μ_{post} en la forma $[\sigma^2/(\sigma^2 + nv^2)] \cdot \mu + [nv^2/(\sigma^2 + nv^2)] \cdot \bar{x}$, como una combinación convexa de las dos. De la expresión anterior se desprende que en el caso de n grande, μ_{post} estará cerca de $\bar{x}$ y no será influida de manera apreciable por la media a priori μ.

Una propiedad de optimalidad de μ_{post} como un estimador de θ se deduce directamente del teorema 3-6. De hecho, se puede probar un resultado más general en este sentido si se emplea este teorema. Suponga que es de interés estimar una función de θ, digamos, $\tau(\theta)$. Para cualquier conjunto de observaciones x de $f(\boldsymbol{x}\,|\,\theta)$, si se define el estadístico $T(\boldsymbol{x})$ como la esperanza a posteriori de $\tau(\theta)$, esto es

$$T(\boldsymbol{x}) = E(\tau(\theta)|\boldsymbol{x}) = \int_{-\infty}^{\infty} \tau(\theta)\pi(\theta|\boldsymbol{x})\,d\theta$$

entonces, del teorema 3-6 se concluye que

$$E[(\tau(\theta) - a(\boldsymbol{x}))^2|\boldsymbol{x}] = \int_{-\infty}^{\infty} (\tau(\theta) - a(\boldsymbol{x}))^2\pi(\theta|\boldsymbol{x})\,d\theta$$

es un mínimo cuando $a(\boldsymbol{x}) = T(\boldsymbol{x})$. En otras palabras, $T(\boldsymbol{x})$ satisface la propiedad

$$E[(\tau(\theta) - T(\boldsymbol{x}))^2|\boldsymbol{x}] = \min_a E[(\tau(\theta) - a(\boldsymbol{x}))^2|\boldsymbol{x}] \text{ por cada } \boldsymbol{x} \tag{12}$$

puesto que $T(\boldsymbol{x})$ es la media de $\tau(\theta)$ con respecto a la densidad a posteriori $\pi(\theta|\boldsymbol{x})$.

En la teoría general de la estimación bayesiana, es común empezar con una función de pérdida $L(\tau(\theta), a)$ que mide la distancia entre el parámetro y un estimador. Luego se busca un estimador, por ejemplo, $\delta^*(\boldsymbol{X})$, con la propiedad de que

$$E[L(\tau(\theta), \delta^*(\boldsymbol{x}))|\boldsymbol{x}] = \min_a E[L(\tau(\theta), a(\boldsymbol{x}))|\boldsymbol{x}] \text{ para cada valor } \boldsymbol{x} \text{ de } \boldsymbol{X} \tag{13}$$

donde la esperanza está en el espacio del parámetro dotado con la densidad a posteriori. Un estimador que satisface la ecuación (13) es llamado *estimador de Bayes de* $\tau(\theta)$ *con respecto a la función de pérdida* $L(\tau(\theta), a)$. En consecuencia, el siguiente teorema es sólo una reformulación de (12):

Teorema 11-7 La media de $\tau(\theta)$ con respecto a la distribución a posteriori $\pi(\theta|\boldsymbol{X})$ es el estimador de Bayes de $\tau(\theta)$ de la función de pérdida del error cuadrático $L(\theta, a) = (\theta - a)^2$.

Otra función de pérdida común es la función de pérdida del error absoluto $L(\theta, a) = |\theta - a|$. En el problema 11.100 se demuestra que la mediana de la densidad a posteriori es el estimador de Bayes de esta función de pérdida.

Teorema 11-8 La mediana de $\tau(\theta)$ con respecto a la distribución a posteriori $\pi(\theta|X)$ es el estimador de Bayes de $\tau(\theta)$ para la función de pérdida del error absoluto $L(\theta, a) = |\theta - a|$.

Cuando $\tau(\theta) = \theta$, estos dos teoremas se reducen a los resultados de optimalidad que se mencionaron en los casos de la media y la mediana a posteriori como estimadores de θ.

EJEMPLO 11.15 Suponga que X es una variable aleatoria binomial con parámetros n y θ, mientras que la densidad a priori de θ es beta con parámetros $\alpha = \beta = 1$. En este caso, los teoremas 11-7 y 11-8 pueden usarse para obtener los estimadores de Bayes de $\theta(1 - \theta)$ para las funciones de pérdida *a*) del error cuadrático y *b*) del error absoluto.

a) La media a posteriori de $\theta(1 - \theta)$ se obtiene a partir del teorema 11-1. Se tiene

$$E(\theta(1 - \theta)|x) = E(\theta|x) - E(\theta^2|x) = \frac{x + 1}{n + 2} - \left[\frac{(x + 1)(x + 2)}{(n + 2)(n + 3)}\right] = \frac{(x + 1)(n - x + 1)}{(n + 2)(n + 3)}$$

b) Es posible obtener numéricamente la mediana de la distribución a posteriori de $\theta(1 - \theta)$ a partir de la distribución a posteriori de θ usando software de computadora. Para demostrar el trabajo implicado se supone que $n = 10$ y $x = 4$. La distribución a posteriori de θ es beta con parámetros 5 y 7. La mediana de $\theta(1 - \theta)$, digamos m, satisface la condición $P(\theta(1 - \theta) \le m) = 0.5$, que es equivalente a la condición de que $P\left(\theta \le \frac{1}{2} - \frac{\sqrt{1 - 4m}}{2}\right) + P\left(\theta \ge \frac{1}{2} + \frac{\sqrt{1 - 4m}}{2}\right) = 0.5$ bajo la distribución beta con parámetros 5 y 7. La solución es $m = 0.247$. (En este caso, la media a posteriori de $\theta(1 - \theta)$ es 0.224.)

ESTIMACIÓN DEL INTERVALO BAYESIANO

Dada la función de densidad a posteriori $\pi(\theta|\boldsymbol{x})$ para un parámetro θ, a cualquier intervalo de θ: $[\theta_L, \theta_U]$ con la propiedad

$$\int_{\theta_L}^{\theta_U} \pi(\theta|x)\, d\theta = 1 - \alpha \tag{14}$$

se le llama *intervalo de credibilidad bayesiano* del $(1 - \alpha) \times 100\%$ para θ. De los diversos intervalos posibles que satisfacen esta propiedad, dos merecen mencionarse: el intervalo de áreas de colas iguales y el intervalo de máxima densidad a posteriori (*MDP*).

El intervalo de áreas de colas iguales del $(1 - \alpha) \times 100\%$ tiene la propiedad de que el área abajo de la densidad a posteriori a la izquierda de θ_L es igual al área a la derecha de θ_U:

$$\int_{-\infty}^{\theta_L} \pi(\theta|x)\, d\theta = \int_{\theta_U}^{\infty} \pi(\theta|x)\, d\theta = (1 - \alpha)/2$$

El requisito del intervalo de *MDP* es que, además de (14), se tiene $\pi(\theta|\boldsymbol{x}) \ge \pi(\theta'|\boldsymbol{x})$ si $\theta \in [\theta_L, \theta_U]$ y $\theta' \notin [\theta_L, \theta_U]$. Es obvio que si $\pi(\theta|\boldsymbol{x})$ no tiene una moda única, entonces es posible que el conjunto de valores de θ que satisface la última condición no sea un intervalo. Para evitar la posibilidad anterior, en este caso se supone que la densidad a posteriori es unimodal. Se concluye de manera directa de esta suposición que $\pi(\theta_L|\boldsymbol{x}) = \pi(\theta_U|\boldsymbol{x})$ y que para cualquier α el intervalo de *MDP* es el más corto de todos los posibles intervalos de credibilidad del $(1 - \alpha) \times 100\%$. Sin embargo, los intervalos de áreas de colas iguales son mucho más fáciles de construir a partir de los percentiles ya disponibles de la mayoría de las distribuciones comunes. Los dos intervalos coinciden cuando la densidad a posteriori es simétrica y unimodal.

EJEMPLO 11.16 Suponga que una muestra aleatoria de tamaño 9 de una distribución normal con media θ desconocida y varianza 1 produce una media muestral de 2.5. También, suponga que la distribución a priori de θ es normal con media 0 y varianza 1. A partir del teorema 11-3 se ve que la distribución a posteriori de θ es normal con media de 2.25 y varianza de 0.1. Un intervalo de credibilidad de colas iguales de 95% para θ está dado por $[\theta_L, \theta_U]$ con θ_L y θ_U iguales, respectivamente, a los percentiles 2.5 y 97.5 de la densidad normal con media de 2.25 y varianza de 0.1. Del apéndice C se deduce que $\theta_L \approx 2.25 - (2.36 \times 0.32) = 1.49$ y $\theta_U \approx 2.25 + (2.36 \times 0.32) = 3.01$. El intervalo bayesiano de credibilidad de colas iguales de 95% (*y* el intervalo de *MDP*, por la simetría de la densidad normal) está dado, en consecuencia, por [1.49, 3.01].

EJEMPLO 11.17 En el problema 6.6 se obtuvieron los intervalos de confianza tradicionales de una media normal θ con base en una muestra de tamaño $n = 200$ bajo el supuesto de que la desviación estándar poblacional era $\sigma = 0.042$. El in-

tervalo de confianza de 95% de la media poblacional resultó ser [0.82, 0.83]. Es aleccionador obtener ahora la probabilidad a posteriori real de este intervalo que se obtuvo suponiendo una distribución a priori normal para θ con media $\mu = 1$ y desviación estándar $v = 0.05$.

A partir del teorema 11-3 se advierte que la densidad a posteriori tiene media $\mu_{\text{post}} \approx 0.825$ y desviación estándar $v_{\text{post}} \approx 0.003$. El área bajo la densidad anterior en el intervalo [0.82, 0.83] es 0.9449.

Cabe mencionar una diferencia básica conceptual entre los intervalos de confianza convencionales y los bayesianos de credibilidad. El enunciado de confianza asociado con un intervalo de confianza de 100 α% para un parámetro θ es el enunciado de probabilidad $P_X(L(\boldsymbol{X}) \leq \theta \leq U(\boldsymbol{X})) = \alpha$ en el espacio muestral de observaciones, con la interpretación de frecuencia de que en muestreos repetidos el intervalo aleatorio $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ encerrará a la constante θ un total de 100 α% de las veces. Sin embargo, dada una muestra aleatoria $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ de observaciones en $\boldsymbol{X}$, el enunciado $P(L(\boldsymbol{x}) \leq \theta \leq U(\boldsymbol{x})) = \alpha$ (en palabras "se tiene una confianza de 100 α% de que θ se ubica entre $L(\boldsymbol{x})$ y $U(\boldsymbol{x})$") no tiene ningún sentido porque sencillamente θ, $L(\boldsymbol{x})$ y $U(\boldsymbol{x})$ son todas constantes.

El enunciado de credibilidad que se asocia con el intervalo bayesiano de credibilidad a 100 α% es el enunciado de probabilidad $P_\Theta(L(\boldsymbol{x}) \leq \theta \leq U(\boldsymbol{x})) = \alpha$ del espacio del parámetro dotado con la densidad de probabilidad $\pi(\theta|\boldsymbol{x})$. Aunque quizás este enunciado no tiene una interpretación de frecuencia, es una descripción de resumen válida de la distribución del parámetro en el sentido en que el intervalo $[L(\boldsymbol{x}), U(\boldsymbol{x})]$ tiene una probabilidad de α bajo la densidad a posteriori $\pi(\theta|\boldsymbol{x})$.

PRUEBAS DE HIPÓTESIS BAYESIANAS

Suponga que se desea probar la hipótesis nula $H_0 : \theta \leq \theta_0$ contra una alternativa $H_1 : \theta > \theta_0$. En este caso, una regla razonable para rechazar H_0 a favor de H_1 puede basarse en la probabilidad a posteriori de la hipótesis nula dados los datos

$$P(H_0|\boldsymbol{x}) = \int_{-\infty}^{\theta_0} \pi(\theta|\boldsymbol{x})\, d\theta \tag{15}$$

Por ejemplo, se puede especificar una $\alpha > 0$ y decidir rechazar H_0 cuando $\boldsymbol{x}$ es tal que $P(H_0|\boldsymbol{x}) \leq \alpha$. Una prueba que se basa en el criterio de rechazo anterior se llama *prueba α de Bayes.*

Nota 3 La probabilidad a posteriori bayesiana de la hipótesis nula que se muestra en (15) es bastante diferente a la prueba de valor *P* (vea la página 215), aunque las dos se confunden a menudo, y la última se refiere con frecuencia y de manera general como la probabilidad de la hipótesis nula.

A continuación se muestra una propiedad de optimalidad que tienen las pruebas α de Bayes. En el capítulo 7 se observó que las cantidades de interés principal cuando se evalúa el desempeño de una prueba son las probabilidades de error tipo I y tipo II para cada θ. Si C es la región crítica de una prueba, entonces estas dos probabilidades están dadas por

$$P_{\text{I}}(\theta) = \begin{cases} \displaystyle\int_C f(\boldsymbol{x}, \theta)\, d\boldsymbol{x}, & \theta \leq \theta_0 \\ 0 & \theta > \theta_0 \end{cases} \quad \text{y} \quad P_{\text{II}}(\theta) = \begin{cases} \displaystyle\int_{C'} f(\boldsymbol{x}, \theta) d\boldsymbol{x}, & \theta > \theta_0 \\ 0 & \theta \leq \theta_0 \end{cases}$$

En el caso de cualquier α especificada, la siguiente media ponderada de estas dos probabilidades se conoce como *riesgo de Bayes* de la prueba.

$$r(C) = (1 - \alpha) \int_{-\infty}^{\theta_0} \int_C \pi(\theta|\boldsymbol{x}) P_{\text{I}}(\theta)\, d\boldsymbol{x}\, d\theta + \alpha \int_{\theta_0}^{\infty} \int_{C'} \pi(\theta|\boldsymbol{x}) P_{\text{II}}(\theta)\, d\boldsymbol{x}\, d\theta \tag{16}$$

Para cada $\boldsymbol{x}$ fija es posible escribir la cantidad del lado derecho como

$$(1 - \alpha)P(\theta \leq \theta_0|\boldsymbol{x})I_C(\boldsymbol{x}) + \alpha P(\theta > \theta_0|\boldsymbol{x})I_{C'}(\boldsymbol{x}) = (1 - \alpha)P(\theta \leq \theta_0|\boldsymbol{x})I_C(\boldsymbol{x}) + \alpha P(\theta > \theta_0|\boldsymbol{x})(1 - I_{C'}(\boldsymbol{x}))$$
$$= \alpha P(\theta > \theta_0|\boldsymbol{x}) + [(1 - \alpha)P(\theta \leq \theta_0|\boldsymbol{x})I_C(\boldsymbol{x}) - \alpha P(\theta > \theta_0|\boldsymbol{x})I_C(\boldsymbol{x})]$$

donde $I_C(\boldsymbol{x})$ denota la función indicadora del conjunto E. El término dentro de los corchetes se minimiza cuando la región crítica C se define de modo que

$$I_C(\boldsymbol{x}) = \begin{cases} 1 & \text{si } (1 - \alpha)P(\theta \leq \theta_0|\boldsymbol{x}) \leq \alpha P(\theta > \theta_0|\boldsymbol{x}) \\ 0 & \text{si no es así} \end{cases}$$

La expresión anterior muestra que $r(C)$ se minimiza cuando C consiste en aquellos puntos datos x para los cuales $P(\theta \leq \theta_0 | \boldsymbol{x}) \leq \alpha$.

En consecuencia, se ha establecido que la *prueba α de Bayes* minimiza el riesgo de Bayes definido por (16). En general, se tiene el siguiente teorema.

Teorema 11-9 En cualquier subconjunto Θ_0 del espacio de parámetros, entre todas las pruebas de la hipótesis nula H_0: $\theta \in \Theta_0$ contra la alternativa H_1: $\theta \in \Theta_0'$, la *prueba α de Bayes*, que rechaza a H_0 si $P(\theta \in \Theta_0 | \boldsymbol{x}) \leq \alpha$, minimiza el riesgo de Bayes definido por

$$r(C) = (1 - \alpha) \int_{\Theta_0} \int_{C} \pi(\theta | \boldsymbol{x}) P_{\mathrm{I}}(\theta)\, d\boldsymbol{x}\, d\theta + \alpha \int_{\Theta_0'} \int_{C'} \pi(\theta | \boldsymbol{x}) P_{\mathrm{II}}(\theta)\, d\boldsymbol{x}\, d\theta$$

EJEMPLO 11.18 Suponga que se sabe que el tiempo de reacción (en segundos) de un individuo a un estímulo particular está distribuido normalmente con media θ desconocida pero desviación estándar conocida de 0.30 seg. La densidad a priori de θ es normal con $\mu = 0.4$ seg y $v^2 = 0.13$. Una muestra de 20 observaciones produjo un tiempo de reacción medio de 0.35 seg. Se desea probar la hipótesis nula $H_0 : \theta \leq 0.3$ contra la alternativa $H_1 : \theta > 0.3$ mediante una prueba 0.05 de Bayes.

Según el teorema 11-3, la densidad a posteriori es normal con media de 0.352 y varianza de 0.004. La probabilidad a posteriori de H_0 está dada por $P(\theta \leq 0.3) = P\left(Z \leq \dfrac{0.3 - 0.352}{0.063}\right) \approx 0.20$. Puesto que la probabilidad anterior es mayor que 0.05, no es posible rechazar H_0.

EJEMPLO 11.19 X es una variable aleatoria de Bernoulli con una probabilidad de éxito θ, que se sabe que es igual a 0.3, o bien 0.6. Se quiere probar la hipótesis nula $H_0 : \theta = 0.3$ contra la alternativa $H_1 : \theta = 0.6$ mediante una prueba 0.05 de Bayes bajo la suposición de una distribución de probabilidad a priori imprecisa para θ: $P(\theta = 0.3) = P(\theta = 0.6) = 0.5$. Una muestra de 30 ensayos en X produce 16 éxitos. Para probar el criterio de rechazo de la prueba 0.05 de Bayes, se necesita la probabilidad a posteriori de la hipótesis nula:

$$P(\theta = 0.3 | x = 16) = \frac{P(x = 16 | \theta = 0.3) \cdot P(\theta = 0.3)}{P(x = 16 | \theta = 0.3) \cdot P(\theta = 0.3) + P(x = 16 | \theta = 0.6) \cdot P(\theta = 0.6)}$$

$$= \frac{(0.0042)(0.5)}{(0.0042)(0.5) + (0.1101)(0.5)} \approx 0.037$$

Como la probabilidad anterior es menor que 0.05 se rechaza la hipótesis nula.

FACTORES DE BAYES

Cuando la distribución a priori que está involucrada es propia, la inferencia estadística bayesiana se puede formular en el lenguaje de posibilidades (vea la página 5) mediante el empleo de los *factores de Bayes,* los cuales es posible considerar como los análogos bayesianos de las razones de posibilidades sobre los cuales se basa la mayoría de las pruebas clásicas del capítulo 7.

Considere el problema de prueba de hipótesis que se presentó en la sección anterior. Es de interés probar la hipótesis nula $H_0 : \theta \in \Theta_0$ contra la alternativa $H_1 : \theta \in \Theta_0'$. Las cantidades

$$\frac{P(H_0)}{P(H_1)} = \frac{\displaystyle\int_{\Theta_0} \pi(\theta)\, d\theta}{\displaystyle\int_{\Theta_0'} \pi(\theta)\, d\theta} \qquad \text{y} \qquad \frac{P(H_0 | \boldsymbol{x})}{P(H_1 | \boldsymbol{x})} = \frac{\displaystyle\int_{\Theta_0} \pi(\theta | \boldsymbol{x})\, d\theta}{\displaystyle\int_{\Theta_0'} \pi(\theta | \boldsymbol{x})\, d\theta} \tag{17}$$

se conocen, respectivamente, como *razones de posibilidades a priori y a posteriori* de H_0 relativa a H_1. El *factor de Bayes* (*FB,* para abreviar) se define como la razón de posibilidades a posteriori sobre la razón de posibilidades a priori. Con base en el hecho de que $\pi(\theta \mid \boldsymbol{x}) \propto f(x \mid \theta)\, \pi(\theta)$ es posible escribir el factor de Bayes de la siguiente manera:

$$FB = \frac{\text{razón de posibilidades a posteriori}}{\text{razón de posibilidaes a priori}} = \left(\frac{P(H_0 | \boldsymbol{x})}{P(H_1 | \boldsymbol{x})}\right) \Big/ \left(\frac{P(H_0)}{P(H_1)}\right) = \frac{\dfrac{1}{P(H_0)} \displaystyle\int_{\Theta_0} f(\boldsymbol{x} | \theta) \pi(\theta)\, d\theta}{\dfrac{1}{P(H_1)} \displaystyle\int_{\Theta_0'} f(\boldsymbol{x} | \theta) \pi(\theta)\, d\theta} \tag{18}$$

En consecuencia, el factor de Bayes es la proporción de los marginales (o los promedios) de la probabilidad bajo las dos hipótesis. Se puede observar también a partir de (18) que cuando ambas hipótesis son sencillas, por ejemplo $H_0 : \theta = \theta_0$ y $H_1 : \theta = \theta_1$ el factor de Bayes se convierte en la familiar razón de posibilidades de la inferencia clásica:

$$FB = \frac{f(\boldsymbol{x}|\theta_0)}{f(\boldsymbol{x}|\theta_1)}.$$

EJEMPLO 11.20 En el ejemplo 11.18, calcule el factor de Bayes de la hipótesis nula $H_0 : \theta \leq 0.3$ contra la alternativa, $H_1 : \theta > 0.3$, con base en (18). Se necesita $P(H_0) = P(\theta \leq 0.3)$, donde θ es la variable aleatoria normal con media de 0.4 y varianza de 0.13. Lo que es igual a $P\left(Z \leq \frac{0.3 - 0.4}{0.36}\right) \approx 0.39$. La probabilidad a posteriori de la hipótesis nula, que está en el ejemplo 11.18, es $P(H_0|\boldsymbol{x}) \approx 0.20$. El factor de Bayes es $\left(\frac{P(H_0|\boldsymbol{x})}{P(H_1|\boldsymbol{x})}\right)\Big/\left(\frac{P(H_0)}{P(H_1)}\right) = \left(\frac{0.20}{0.80}\right)\Big/\left(\frac{0.39}{0.61}\right) \approx 0.39$.

EJEMPLO 11.21 Una caja contiene una moneda no cargada y dos cargadas (cada una con $P(\text{“caras”}) = 0.2$). Se toma al azar una moneda de la caja y se lanza 10 veces. Si se obtienen 4 caras, ¿cuál es el factor de Bayes de la hipótesis nula H_0 de que la moneda elegida no está cargada con respecto a la hipótesis alternativa H_1 de que sí está cargada? Las probabilidades a priori son $P(H_0) = 1/3$ y $P(H_1) = 2/3$, por lo que la razón de posibilidades a priori es 0.5. Las probabilidades a posteriori son $P(H_0|\boldsymbol{x}) = \frac{(0.5)^{10}}{(0.5)^{10} + 2(0.2)^4(0.8)^6} \approx 0.54$ y $P(H_1|\boldsymbol{x}) = \frac{(0.2)^4(0.8)^6}{(0.5)^{10} + (0.2)^4(0.8)^6} \approx 0.46$, por lo que la razón de posibilidades a posteriori es $0.54/0.56 \approx 1.16$. En consecuencia, el factor de Bayes es $1.16/0.5 \approx 3.32$. Es posible también conseguir el mismo resultado directamente como la razón de las probabilidades bajo las dos hipótesis $P(\boldsymbol{x}|H_0) = \binom{10}{4}(0.5)^{10}$ y $P(\boldsymbol{x}|H_1) = \binom{10}{4}(0.2)^4(0.8)^6$.

De (18) se deduce que el factor de Bayes cuantifica la fuerza de la evidencia proporcionada por los datos a favor o no de la hipótesis nula relativa a la alternativa. En términos generales, es posible decir que si el factor de Bayes es mayor que 1, los datos observados confirman la hipótesis nula, y si es menor que 1, no la confirman. Además, cuanto mayor sea el factor de Bayes, tanto más fuerte será la evidencia a favor de la hipótesis nula. La calibración del factor de Bayes para reflejar la fuerza real de la evidencia a favor o no de la hipótesis nula es un tema que no se analiza aquí. Sin embargo, es posible probar el siguiente teorema:

Teorema 11-10 La prueba α de Bayes es equivalente a la prueba que rechaza la hipótesis nula si

$$FB \leq \frac{\alpha[1 - P(H_0)]}{(1 - \alpha)P(H_0)}.$$

Para apreciar lo anterior observe que el criterio de rechazo de una prueba α de Bayes, a saber $P(H_0|\boldsymbol{x}) \leq \alpha$, es equivalente a la condición $\frac{P(H_0|\boldsymbol{x})}{P(H_1|\boldsymbol{x})} \leq \frac{\alpha}{1 - \alpha}$ y que esta desigualdad equivale a la condición $FB \leq \frac{\alpha[1 - P(H_0)]}{(1 - \alpha)P(H_0)}$.

Nota 4 Una regla *ad hoc* que en ocasiones se utiliza es rechazar la hipótesis nula si $FB < 1$. Se puede demostrar que dicha regla es equivalente a la prueba α de Bayes con $\alpha = P(H_0)$: rechazar H_0 si $P(H_0|\boldsymbol{x}) \leq P(H_0)$.

EJEMPLO 11.22 Determine el criterio de rechazo en términos del factor de Bayes para la prueba que se usa en el ejemplo 11.19. Se tiene $\alpha = 0.05$ y $P(H_0) = P(\theta = 0.3) = 0.5$. Por tanto, según el teorema 11-10, el criterio de prueba es rechazar H_0 si $FB \leq \frac{(0.05)(0.5)}{(0.95)(0.5)} \approx 0.053$. El factor de Bayes correspondiente a 16 éxitos de 30 ensayos es $\left(\frac{P(H_0|\boldsymbol{x})}{P(H_1|\boldsymbol{x})}\right)\Big/\left(\frac{P(H_0)}{P(H_1)}\right) = \left(\frac{0.037}{1 - 0.037}\right)\Big/\left(\frac{0.5}{0.5}\right) \approx 0.038$, y puesto que es menor que 0.053, se rechaza la hipótesis nula.

EJEMPLO 11.23 En el ejemplo 11.18 suponga que se desea emplear la regla de decisión que rechaza H_0 si el factor de Bayes es menor que 1. Ya se sabe que la probabilidad de la hipótesis nula según la densidad a posteriori de θ es 0.20. En consecuencia, las posibilidades a posteriori de H_0 son $\frac{0.20}{0.80} = \frac{1}{4}$. La probabilidad a priori de H_0 está dada por $P(\theta \leq 0.3) = P\left(Z \leq \frac{0.3 - 0.4}{0.36}\right) \approx 0.39$, porque las posibilidades a priori de H_0 son $\frac{0.39}{0.61}$ El factor de Bayes de H_0 es $(1/4)\ (39/61) \approx 0.39 < 1$. La decisión, en consecuencia, es rechazar H_0.

DISTRIBUCIONES PREDICTIVAS BAYESIANAS

El marco bayesiano permite obtener la distribución condicional de las observaciones futuras con base en una distribución a priori o a posteriori disponible para el parámetro poblacional, que se denominan *distribuciones predictivas*, cuyo proceso fundamental de deducción es la marginalización directa de la distribución conjunta de las observaciones futuras y el parámetro (vea las páginas 40-41).

Suponga que en n ensayos de Bernoulli con una probabilidad de éxito θ desconocida se obtienen x éxitos y que la densidad a priori de θ es beta con parámetros α y β. Si se consideran m ensayos adicionales en la misma población Bernoulli, ¿qué se puede decir acerca del número de éxitos que se obtiene? Se sabe del teorema 11-1 que la distribución a posteriori de θ, dada x, es beta con parámetros $x + \alpha$ y $n - x + \beta$. Si $f(y \mid \theta)$ es la función de probabilidad del número Y de éxitos en los ensayos futuros m, la densidad conjunta de Y y θ es

$$f(y, \theta \mid x) = f(y \mid \theta)\pi(\theta \mid x) = \binom{m}{y}\theta^y(1-\theta)^{m-y} \cdot \frac{\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}}{B(x+\alpha, n-x+\beta)}$$

$$= \binom{m}{y}\frac{\theta^{x+y+\alpha-1}(1-\theta)^{m+n-x-y+\beta-1}}{B(x+\alpha, n-x+\beta)} \qquad 0 < \theta < 1, y = 0, 1, \ldots, m$$

La función de probabilidad predictiva de Y denotada por $f^*(y)$ es la densidad marginal de Y que se obtuvo de la densidad conjunta anterior al integrar respecto a θ:

$$f^*(y) = \int_0^1 \binom{m}{y}\frac{\theta^{x+y+\alpha-1}(1-\theta)^{n+m-x-y+\beta-1}}{B(x+\alpha, n-x+\beta)}d\theta \tag{19}$$

$$= \binom{m}{y}\frac{B(x+y+\alpha, m+n-x-y+\beta)}{B(x+\alpha, n-x+\beta)} \qquad y = 0, 1, \ldots, m \tag{20}$$

Por tanto, se puede enunciar el siguiente teorema:

Teorema 11-11 Si en n ensayos de Bernoulli con probabilidad de éxito desconocida θ se obtienen x éxitos, y la densidad a priori de θ es beta con parámetros α y β, entonces la densidad predictiva del número de éxitos Y en los m ensayos futuros en la misma población de Bernoulli está dada por (20).

Nota 5 A partir de (19) es evidente que es posible considerar a $f^*(y)$ también como la esperanza, $E_\theta(f(y \mid \theta))$, de la función de probabilidad de Y con respecto a la densidad a posteriori $\pi(\theta \mid x)$ de θ.

EJEMPLO 11.24 Suponga que se obtuvieron 7 éxitos en 10 ensayos de Bernoulli con la probabilidad de éxito θ. Se desea realizar un conjunto independiente de 8 ensayos de Bernoulli más con la misma probabilidad de éxito. ¿Qué se puede decir acerca del número de éxitos futuros si θ tiene una densidad a priori uniforme en el intervalo [0, 1]?

La distribución predictiva del número de éxitos futuros puede obtenerse de (20) con $\alpha = \beta = 1, n = 10, m = 8$ y $x = 7$:

$$f^*(y) = \binom{8}{y}\frac{B(y+8, 12-y)}{B(8, 4)} \qquad y = 0, 1, \ldots, 8$$

La tabla 11-5 resume los resultados numéricos.

Tabla 11-5

y	0	1	2	3	4	5	6	7	8
$f^*(y)$	0.002	0.012	0.040	0.089	0.153	0.210	0.227	0.182	0.085

Nota 6 En una nota anterior después del teorema 11-5 sobre la *ley de sucesión de Laplace* se señaló que si todos los n ensayos de un experimento binomial resultaran en éxitos, sería posible calcular la probabilidad de que un ensayo futuro también resulte en un éxito a partir de la media a posteriori del parámetro de éxito θ, a saber $(n+1)/(n+2)$. El mismo resultado puede obtenerse como un caso especial de (20) con $\alpha = \beta = 1$, $m = 1$ y $x = n$. La distribución predictiva de una observación futura resulta ser binomial con probabilidad de éxito $(n+1)/(n+2)$. Sin embargo, los dos enfoques no llevan a los mismos resultados más allá de $n = 1$. Por ejemplo, si se considera la media a posteriori $(n+1)/(n+2)$ como la probabilidad de éxito para m ensayos futuros, la probabilidad de que todos sean éxitos sería $[(n+1)/(n+2)]^m$, pero (20) da $(n+1)/(m+n+1)$.

EJEMPLO 11.25 En el ejemplo 11.24 suponga que es de interés predecir el resultado de los primeros 10 ensayos de Bernoulli antes de que se ejecuten. Determine la distribución predictiva del número de éxitos, digamos X, en los 10 ensayos, de nuevo bajo el supuesto de que θ tiene una densidad a priori uniforme en el intervalo [0,1].

La distribución conjunta de X y θ está dada por

$$f(x;\theta) = \binom{10}{x}\theta^x(1-\theta)^{10-x}\cdot 1 \qquad 0<\theta<1 \qquad x = 0, 1, \ldots, 10$$

Es posible obtener la densidad marginal de x a partir de lo anterior si se integra respecto de θ:

$$f^*(x) = \int_0^1 \binom{10}{x}\theta^x(1-\theta)^{10-x}d\theta = \binom{10}{x}B(x+1, 11-x) = \frac{1}{11} \qquad y = 0, 1, \ldots, 10$$

Nota 7 Las distribuciones predictivas que se obtuvieron en los ejemplos 11.24 y 11.25 difieren en el sentido en que se basan, respectivamente, en distribuciones a posteriori y a priori del parámetro. Con frecuencia se establece una distinción entre las distribuciones predictivas a priori y a posteriori para indicar la naturaleza de la distribución del parámetro que se utiliza.

Las distribuciones predictivas de las muestras normales futuras se pueden obtener de manera análoga. En el teorema 11-3 se observó que si se tiene una muestra de tamaño n de una distribución normal con media θ desconocida y varianza σ^2 conocida, y si θ es normal con media μ y varianza v^2, entonces la distribución a posteriori de θ también es normal, con media μ_{post} y varianza v^2_{post} dadas por (7). Suponga que se hace otra observación, por ejemplo Y, en la población original. A continuación se demuestra que la distribución predictiva de Y es normal con media μ_{post} y varianza $\sigma^2 + v^2_{\text{post}}$.

La densidad predictiva $f^*(y)$ de Y está dada por

$$f^*(y) = f(y|\boldsymbol{x}) = \int f(y,\theta|\boldsymbol{x})d\theta = \int f(y|\theta)\cdot\pi(\theta|\boldsymbol{x})\,d\theta$$

$$\propto \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(y-\theta)^2}e^{-\frac{1}{2v^2_{\text{post}}}(\theta-\mu_{\text{post}})^2}d\theta$$

Después de simplificar se tiene

$$f^*(y) = \int_{-\infty}^{\infty} e^{-\frac{1}{2(\sigma^2v^2_{\text{post}})/(\sigma^2+v^2_{\text{post}})}\left[\theta-\frac{(v^2_{\text{post}}y+\sigma^2\mu_{\text{post}})}{(\sigma^2+v^2_{\text{post}})}\right]^2} \times e^{-\frac{1}{2(\sigma^2v^2_{\text{post}})/(\sigma^2+v^2_{\text{post}})}\left[-\left(\frac{v^2_{\text{post}}y+\sigma^2\mu_{\text{post}}}{\sigma^2+v^2_{\text{post}}}\right)^2+\frac{v^2_{\text{post}}y+\sigma^2\mu_{\text{post}}}{\sigma^2+v^2_{\text{post}}}\right]}d\theta$$

Es factible simplificar aún más el exponente del segundo factor para obtener

$$f^*(y) \propto \int_{-\infty}^{\infty} e^{-\frac{1}{2(\sigma^2v^2_{\text{post}})/(\sigma^2+v^2_{\text{post}})}\left[\theta-\frac{(v^2_{\text{post}}y+\sigma^2\mu_{\text{post}})}{\sigma^2+v^2_{\text{post}}}\right]^2} \times e^{-\frac{1}{2(\sigma^2+v^2_{\text{post}})}(y-\mu_{\text{post}})^2}d\theta$$

En este caso, el segundo factor es independiente de θ. El primer factor es una densidad normal en θ y se integra en una expresión libre de θ y y. Por tanto, la integral anterior se convierte en

$$e^{-\frac{1}{2(\sigma^2+v^2_{\text{post}})}(y-\mu_{\text{post}})^2}$$

Esta expresión puede reconocerse como una densidad normal con media μ_{post} y varianza $\sigma^2 + v^2_{\text{post}}$. En consecuencia, se ve que la densidad predictiva de una observación futura Y es normal con media igual a la media a posteriori de θ y varianza igual a la suma de la varianza poblacional y la varianza a posteriori de θ.

El siguiente teorema es una generalización directa del resultado anterior (vea el problema 11.96).

Teorema 11-12 Suponga que una muestra aleatoria de tamaño n se toma de una distribución normal con media θ desconocida y varianza σ^2 conocida y que la distribución a priori de θ es normal con media μ y varianza v^2. Si una segunda muestra independiente de tamaño m se toma de la misma población, entonces la distribución predictiva de la media muestral es normal con media μ_{post} y varianza $\left(\frac{\sigma^2}{m} + v^2_{\text{post}}\right)$, donde $\mu_{\text{post}} = \frac{\sigma^2\mu + nv^2\bar{x}}{\sigma^2 + nv^2}$, $v^2_{\text{post}} = \frac{\sigma^2v^2}{\sigma^2 + nv^2}$, y $\bar{x}$ es la media de la primera muestra, de tamaño n.

EJEMPLO 11.26 El peso de los paquetes que envía una compañía sigue una distribución normal con media de θ libras y varianza de 8. Si los primeros 25 paquetes enviados en un día tienen un peso promedio de 15 libras, ¿cuáles son las probabilidades de que los siguientes 25 paquetes excedan un peso promedio de 16 libras? Suponga que θ tiene una distribución a priori normal con media de 12 y varianza de 9.

A partir del teorema 11-12, la media y la varianza de la densidad predictiva del peso promedio de la futura muestra están dadas por 14.90 y 0.31. La probabilidad que se necesita es $P(\overline{Y} > 16) = P(Z > 1.98) = 0.0234$, por lo que las posibilidades son alrededor de 2% de que el peso promedio futuro sea mayor a 16 libras.

Los resúmenes puntual y de intervalo de la densidad predictiva se pueden obtener igual que en el caso de la densidad a posteriori de un parámetro, y sirven para propósitos similares. Por ejemplo, dada la función de densidad predictiva $f^*(\bar{y})$ de la media muestral $\overline{Y}$ de una futura muestra de una población, se pueden usar la esperanza, la mediana o la moda de $f^*(\bar{y})$ como una *estimación puntual predictiva* de $\overline{Y}$. También es posible usar los intervalos $[\bar{y}_L, \bar{y}_U]$ que satisfacen la propiedad

$$\int_{\bar{y}_L}^{\bar{y}_U} f^*(\bar{y})\, d\bar{y} = 1 - \alpha \tag{21}$$

como *intervalos predictivos* bayesianos del $(1 - \alpha) \times 100\%$ para $\overline{Y}$, y se pueden definir los intervalos predictivos de áreas de colas iguales y los de *MDP* como en el caso de los intervalos de credibilidad de un parámetro.

EJEMPLO 11.27 En el ejemplo 11.24, estime la *a*) media, *b*) mediana y *c*) moda predictivas del número de los éxitos futuros.

a) La distribución predictiva de Y está dada en la tabla 11-5. El número de la media predictiva de los éxitos futuros es la esperanza de Y, la cual es 5.34.
b) La mediana predictiva está entre 5 y 6, y se puede tomar como 5.5.
c) La moda predictiva es 6.

EJEMPLO 11.28 En el ejemplo 11.26 determine el intervalo predictivo del área de colas iguales de 95% para el peso promedio de los 25 paquetes futuros.

La distribución predictiva es normal con media de 14.90 y varianza de 0.31. El intervalo de colas iguales de 95% está dado por $14.90 \pm (1.96 \times 0.56) = [13.8, 16.0]$

PROBLEMAS RESUELTOS

PROBABILIDAD SUBJETIVA

11.1. Identificar el tipo de la probabilidad que se emplea en los siguientes casos: *a*) La probabilidad de que mi hija asistirá a la universidad es 0.9. *b*) Las posibilidades de conseguir tres caras en tres lanzamientos de una moneda son 1 en 8. *c*) Tengo una confianza de 40% de que llueva el 4 de julio de este año porque sucedió 12 veces en los últimos 30 años. *d*) Confío un 70% de que la varianza de esta distribución no excede 3.5. *e*) Algunos economistas creen que hay una probabilidad mayor al 50% de que la economía entrará en una recesión el próximo año. *f*) Las posibilidades son sólo de 2% de que ella fallará los dos tiros libres. *g*) Tengo una confianza de 90% de que esta moneda esté cargada. *h*) La probabilidad de que todos los niños sean varones en una familia con tres hijos es de alrededor de 0.11. *i*) Las probabilidades son 3 en 1 de que los Badgers no lleguen al Súper Tazón este año. *j*) Usted tiene una posibilidad en un millón de ganar esta lotería. *k*) Usted tiene una posibilidad de más de 50% de encontrar una tienda que maneje este producto.

a), *d*), *e*), *g*), *i*): subjetiva; *b*), *j*), *k*): clásica; *c*), *f*), *h*): frecuencia.

DISTRIBUCIONES A PRIORI Y A POSTERIORI

11.2. Una caja contiene dos monedas no cargadas y dos cargadas [cada una con $P(\text{“caras”}) = 0.3$]. Se elige una moneda al azar de la caja y se lanza cuatro veces. Si se obtienen dos caras y dos cruces, estimar las probabilidades a posteriori del evento F de que la moneda seleccionada no esté cargada y el evento B de que la moneda sí lo esté.

Consideremos que D denota al evento (datos) en el que se obtienen dos caras y dos cruces en cuatro lanzamientos. Del teorema de Bayes se deduce

$$P(F|D) = \frac{P(D|F)P(F)}{P(D|F)P(F) + P(D|B)P(B)} = \frac{\left[\binom{4}{2}(0.5)^4\right] \cdot \left(\frac{1}{2}\right)}{\left[\binom{4}{2}(0.5)^4\right] \cdot \left(\frac{1}{2}\right) + \left[\binom{4}{2}(0.3)^2(0.7)^2\right] \cdot \left(\frac{1}{2}\right)} = \frac{625}{1066} \approx 0.59$$

$$P(B|D) = 1 - P(F|D) = 441/1066 \approx 0.41$$

11.3. Verificar los valores de la probabilidad a posteriori dados en la tabla 11-2.

$$P(\theta = 0.2|D) = \frac{P(D|\theta = 0.2)P(\theta = 0.2)}{P(D|\theta = 0.2)P(\theta = 0.2) + P(D|\theta = 0.5)P(\theta = 0.5)} = \frac{[3(0.2)^2(0.8)] \cdot \left(\frac{1}{2}\right)}{[3(0.2)^2(0.8)] \cdot \left(\frac{1}{2}\right) + [3(0.5)^3] \cdot \left(\frac{1}{2}\right)}$$

$$= \frac{32}{157} \approx 0.20$$

$$P(\theta = 0.5|D) = 1 - P(\theta = 0.2|D) = \frac{125}{157} \approx 0.80$$

11.4. La variable aleatoria X tiene una distribución de Poisson con un parámetro desconocido λ. Se determinó que λ tiene la función de probabilidad a priori subjetiva dada en la tabla 11-6. Una muestra aleatoria de tamaño 2 produce los valores 2 y 0 de X. Calcular la distribución a posteriori de λ.

Tabla 11-6

λ	0.5	1.0	1.5
$\pi(\lambda)$	1/2	1/3	1/6

La probabilidad de los datos es $f(\boldsymbol{x}|\lambda) = e^{-2\lambda}\frac{\lambda^{x_1+x_2}}{x_1!x_2!}$. La densidad a posteriori es (en el caso de factores independientes de λ)

$$\pi(\lambda|\boldsymbol{x}) \propto e^{-2\lambda}\lambda^{x_1+x_2}\pi(\lambda) \propto e^{-2\lambda}\lambda^2\pi(\lambda) \qquad \text{para} \qquad \lambda = 0.5, 1, 1.5$$

Los resultados se muestran en la tabla 11-7.

Tabla 11-7

λ	0.5	1.0	1.5
$\pi(\lambda|\boldsymbol{x})$	0.42	0.41	0.17

11.5. En un lote de n pernos que produjo una máquina, un número desconocido ρ es defectuoso. Suponer que ρ tiene una distribución a priori binomial con parámetro p. Determinar la distribución a posteriori de ρ si se elige un perno que al azar del lote sea *a*) defectuoso y *b*) no defectuoso.

a) Se da la función de probabilidad a priori $\pi(\rho) = \binom{n}{\rho}p^\rho(1-p)^{n-\rho}$, $\rho = 0, 1, \ldots, n$. La función de probabilidad a posteriori de ρ, dado el evento D = "defectuoso", es $\pi(\rho|D) \propto \frac{\rho}{n} \cdot \binom{n}{\rho}p^\rho(1-p)^{n-\rho}$, $\rho = 0,1, \ldots, n$.

$$= \binom{n-1}{\rho-1}p^\rho(1-p)^{n-\rho}, \rho = 1, \ldots, n$$

Puesto que $\sum_{\rho=1}^{n}\binom{n-1}{\rho-1}p^{\rho-1}(1-p)^{n-\rho} = 1$, la constante de proporcionalidad en la función de la probabilidad anterior debe ser $\frac{1}{p}$. Por tanto, $\pi(\rho|D) = \binom{n-1}{\rho-1}p^{\rho-1}(1-p)^{n-\rho}$, $\rho = 1, \ldots, n$.

b) $\pi(\rho|D') \propto \dfrac{n-\rho}{n} \cdot \binom{n}{\rho} p^{\rho}(1-p)^{n-\rho}, \rho = 0, 1, \ldots, n-1$

$$= \binom{n-1}{\rho} p^{\rho}(1-p)^{n-\rho}, \rho = 0, \ldots, n-1$$

Puesto que $\sum_{\rho=0}^{n-1} \binom{n-1}{\rho} p^{\rho}(1-p)^{n-1-\rho} = 1$, la constante de proporcionalidad en la función de probabilidad anterior debe ser $\dfrac{1}{1-p}$. En consecuencia, $\pi(\rho|D') = \binom{n-1}{\rho} p^{\rho}(1-p)^{n-1-\rho}, \rho = 0, \ldots, n-1$.

11.6. X es una variable aleatoria binomial con n conocido y una probabilidad de éxito θ desconocida. Determinar la densidad a posteriori de θ bajo el supuesto de una densidad a priori $\pi(\theta)$ igual a *a)* 2θ, $0 < \theta < 1$; *b)* $3\theta^2$, $0 < \theta < 1$; *c)* $4\theta^3$, $0 < \theta < 1$.

a) $\pi(\theta|\boldsymbol{x}) \propto \theta^x(1-\theta)^{n-x} \cdot \theta = \theta^{x+1}(1-\theta)^{n-x}$, $0 < \theta < 1$.

Puesto que ésta es una densidad beta con parámetros $x+2$ y $n-x+1$, la constante de normalización es $1/B(x+2, n-x+1)$ y se obtiene $\pi(\theta|\boldsymbol{x}) = \dfrac{1}{B(x+2, n-x+1)}\theta^{x+1}(1-\theta)^{n-x}$, $0 < \theta < 1$.

b) La a posteriori es la densidad beta $\pi(\theta|\boldsymbol{x}) = \dfrac{1}{B(x+3, n-x+1)}\theta^{x+2}(1-\theta)^{n-x}$, $0 < \theta < 1$.

c) La a posteriori es la densidad beta $\pi(\theta|\boldsymbol{x}) = \dfrac{1}{B(x+4, n-x+1)}\theta^{x+3}(1-\theta)^{n-x}$, $0 < \theta < 1$.

11.7. Se toma una muestra aleatoria $x = (x_1, x_2, \ldots, x_n)$ de tamaño n de una población con una función de densidad $f(x|\theta) = 3\theta x^2 e^{-\theta x^3}$, $x > 0$. θ tiene una densidad gamma a priori con parámetros α y β. Determinar la densidad a posteriori de θ.

$\pi(\theta|\boldsymbol{x}) \propto \theta^n e^{-\theta\sum x^3} \cdot \theta^{\alpha-1}e^{-\theta/\beta} \propto \theta^{n+\alpha-1}e^{-\theta\left(\sum x^3 + \frac{1}{\beta}\right)}$. Ésta se puede reconocer como una densidad gamma con parámetros $n+\alpha$ y $\dfrac{\beta}{1+\beta\sum x^3}$. En consecuencia, la constante de normalización debería ser $\left(\dfrac{1+\beta\sum x^3}{\beta}\right)^{n+\alpha} \cdot \dfrac{1}{\Gamma(n+\alpha)}$ y la densidad a posteriori es $\pi(\theta|\boldsymbol{x}) = \dfrac{1}{\Gamma(n+\alpha)}\left(\dfrac{1+\beta\sum x^3}{\beta}\right)^{n+\alpha}\theta^{n+\alpha-1}e^{-\theta\left(\sum x^3 + \frac{1}{\beta}\right)}$, $\theta > 0$.

11.8. X es normal con media 0 y precisión desconocida ξ que tiene una densidad gamma a priori con parámetros α y β. Determinar la distribución a posteriori de ξ con base en la muestra aleatoria $x = (x_1, x_2, \ldots, x_n)$ de X.

$\pi(\xi|\boldsymbol{x}) \propto \xi^{n/2}e^{-\frac{\xi}{2}\sum x^2} \cdot \xi^{\alpha-1}e^{-\xi/\beta} \propto \xi^{\frac{n}{2}+\alpha-1}e^{-\xi\left(\frac{\sum x^2}{2} + \frac{1}{\beta}\right)}$, $\xi > 0$

En consecuencia, ξ tiene una distribución gamma con parámetros $\dfrac{n}{2} + \alpha$ y $\dfrac{2\beta}{\beta\sum x^2 + 2}$.

MUESTREO DE UNA POBLACIÓN BINOMIAL

11.9. Una encuesta entre 100 votantes elegidos de manera aleatoria de todo el padrón de un distrito dado indicó que 55% de ellos estuvieron a favor de un candidato. Suponer que se cree que antes de la encuesta la proporción verdadera θ de votantes de ese distrito a favor del candidato tiene una densidad uniforme sobre el intervalo [0,1]. Determinar la densidad a posteriori de θ.

Con base en el teorema 11-1 con $n = 100$ y $x = 55$, la densidad a posteriori de θ es beta con parámetros $\alpha = 56$ y $\beta = 46$.

11.10. En 40 lanzamientos de una moneda se obtuvieron 24 caras. Determinar la distribución a posteriori de la proporción θ de caras que se obtendría en un número ilimitado de lanzamientos de la moneda. Utilizar una a priori uniforme para θ.

Según el teorema 11-1, la densidad a posteriori de θ es beta con $\alpha = 25$ y $\beta = 17$.

11.11. Una encuesta para predecir el resultado de un próximo referéndum determinó que 480 de 1000 personas encuestadas estuvieron a favor de él. ¿Cuáles son las probabilidades de que el referéndum no se apruebe?

Suponemos una distribución a priori imprecisa (uniforme en [0,1]) para la proporción θ de personas que está a favor del referéndum. La distribución a posteriori de θ, dados los resultados de la encuesta, es beta con paráme-

tros 481, 521. Se necesita la probabilidad de que $\theta < 0.5$. Software de computadora proporciona 0.90 para esta probabilidad, por lo que es posible tener una confianza de 90% de que el referéndum no se aprobará.

11.12. En el problema anterior, suponer que se encuestan 1000 personas adicionales, 530 de las cuales estuvieron a favor del referéndum. ¿Qué es lo que se puede concluir en este caso?

Ahora se considera que la a priori es beta con parámetros 481 y 521. La a posteriori se convierte en beta con parámetros 1011 y 991. La probabilidad para $\theta < 0.5$ es 0.33, lo que significa que existe una probabilidad ahora de sólo 33% de que no se apruebe el referéndum.

MUESTREO DE UNA POBLACIÓN DE POISSON

11.13. El número de accidentes durante un periodo de seis meses en una intersección tiene una distribución de Poisson con media λ. Se cree que λ tiene una densidad a priori gamma con parámetros $\alpha = 2$ y $\beta = 5$. Si se observaron ocho accidentes durante los primeros seis meses del año, calcular *a*) la densidad a posteriori, *b*) la media a posteriori y *c*) la varianza a posteriori.

a) Por el teorema 11-2 sabemos que la densidad a posteriori es gamma con parámetros $n\bar{x} + \alpha = 10$ y $\beta/(1 + n\beta) = 5/6$.

b) A partir de (32) del capítulo 4, la media a posteriori es $50/6 \approx 8.33$ y *c*) la varianza a posteriori es $= 250/36 \approx 6.94$.

11.14. El número de defectos en un carrete de hilo de 1000 pies que fabrica una máquina tiene una distribución de Poisson con media λ desconocida. La distribución a priori de λ es gamma con parámetros $\alpha = 2$ y $\beta = 1$. Se encontró un total de 23 defectos en una muestra de 10 carretes. Determinar la densidad a posteriori de λ.

De acuerdo con el teorema 11-2, la densidad a posteriori es gamma con parámetros $n\bar{x} + \alpha = 23 + 2 = 25$ y $\beta/(1 + n\beta) = 1/11 \approx 0.091$.

MUESTREO A PARTIR DE UNA POBLACIÓN NORMAL

11.15. En una muestra de 100 mediciones del diámetro de una esfera se obtuvo una media de $\bar{x} = 4.38$ pulgadas. Con base en experiencias previas se sabe que el diámetro está distribuido de manera normal con media θ desconocida y varianza de 0.36. Determinar la densidad a posteriori de θ bajo el supuesto de una densidad a priori normal con media de 4.5 pulgadas y varianza de 0.4.

Con base en el teorema 11-3 se observa que la densidad a posteriori es normal con media de 4.381 y varianza de 0.004.

11.16. Se sabe que el tiempo de reacción de un individuo a cierto estímulo tiene una distribución normal con media θ desconocida pero una desviación estándar conocida de 0.35 segundos. Una muestra de 20 observaciones produce un tiempo de reacción medio de 1.18 segundos. Suponer que la densidad a priori de θ es normal con media $\mu = 1$ segundos y varianza $v^2 = 0.13$. Determinar la densidad a posteriori de θ.

Según el teorema 11-3, la densidad a posteriori es normal con media de 1.17 y varianza de 0.006.

11.17. Se toma una muestra aleatoria de 25 observaciones de una población normal con media θ desconocida y varianza igual a 16. La distribución a priori de θ es normal estándar. Calcular *a*) la media a posteriori y *b*) su precisión. *c*) Determinar la precisión del estimador de máxima verosimilitud.

a) Según el teorema 11-3, la media a posteriori de θ es $\left(\dfrac{25}{25 + 16}\right)\bar{x} = \dfrac{25\bar{x}}{41}$.

b) La precisión del estimador es el recíproco de su varianza. La varianza de $\dfrac{25\bar{x}}{41} = \left(\dfrac{25}{41}\right)^2 \dfrac{16}{25} = 0.24$, por lo que la precisión es de aproximadamente 4.2.

c) El estimador de máxima verosimilitud de θ es $\bar{x}$. Su varianza corresponde a $16/25$, de modo que la precisión es alrededor de 1.6.

11.18. X es normal con media 0 y precisión no conocida ξ, la cual tiene una distribución a priori gamma con parámetros α y β. Determinar la distribución a posteriori de ξ con base en una muestra aleatoria $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ de X.

$$\pi(\xi \mid \boldsymbol{x}) \propto \xi^{n/2} e^{-\frac{\xi}{2}\sum x^2} \cdot \xi^{\alpha-1} e^{-\xi/\beta} = \xi^{\frac{n}{2}+\alpha-1} e^{-\xi\left(\sum x^2 + \frac{1}{\beta}\right)}$$

Por tanto, ξ tiene una distribución gamma con parámetros $\dfrac{n}{2} + \alpha$ y $\dfrac{\beta}{\beta\sum x^2 + 1}$.

DISTRIBUCIONES A PRIORI IMPROPIAS

11.19. Una densidad a priori impropia para una media λ de Poisson se define por medio de $\pi(\lambda) = 1$, $\lambda > 0$. Demostrar que la densidad a posteriori en este caso es gamma con parámetros $n\bar{x} + 1$ y $\frac{1}{n}$.

Considerando el vector de observación $\boldsymbol{x}$, la densidad a posteriori de λ es $\pi(\lambda|x) \propto e^{-n\lambda}\lambda^{n\bar{x}}$. El resultado se infiere basado en que la densidad anterior es de forma gamma con parámetros $n\bar{x} + 1$ y $\frac{1}{n}$.

11.20. Otra densidad a priori impropia para la media λ de Poisson es $\pi(\lambda) = 1/\lambda$, $\lambda > 0$. Demostrar que la densidad a posteriori en este caso es gamma.

Se tiene $\pi(\lambda|x) \propto e^{-n\lambda}\lambda^{n\bar{x}} \cdot \frac{1}{\lambda} \propto e^{-n\lambda}\lambda^{n\bar{x}-1}$. La a posteriori es, por tanto, gamma con parámetros $n\bar{x}$ y $\frac{1}{n}$.

11.21. Una densidad a priori impropia para la media λ de Poisson, la cual se conoce como *a priori de Jeffrey* para la media de Poisson, está dada por $\pi(\lambda) = 1/\sqrt{\lambda}$, $\lambda > 0$. Determinar la densidad a posteriori bajo esta a priori.

Si se considera el vector de observación $\boldsymbol{x}$, la densidad a posteriori de λ es $\pi(\lambda|\boldsymbol{x}) \propto e^{-n\lambda}\lambda^{n\bar{x}} \cdot \left(\frac{1}{\sqrt{\lambda}}\right) \propto e^{-n\lambda}\lambda^{n\bar{x}-\frac{1}{2}}$, $\lambda > 0$.

Ésta es una densidad gamma con parámetros $n\bar{x} + \frac{1}{2}$ y $\frac{1}{n}$.

11.22. X es una binomial con n conocida y probabilidad de éxito θ no conocida. Una densidad a priori impropia para θ, llamada *a priori de Haldane*, está dada por $\pi(\theta) = \frac{1}{\theta(1-\theta)}$, $0 < \theta < 1$ Determinar la densidad a posteriori de θ basada en la observación x.

$\pi(\theta|x) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \cdot \frac{1}{\theta(1-\theta)} = \binom{n}{x}\theta^{x-1}(1-\theta)^{n-x-1}$, $0 < \theta < 1$, la cual es una densidad beta con $\alpha = x$ y $\beta = n - x$, así que se tiene a $\pi(\theta|x) = \frac{\theta^{x-1}(1-\theta)^{n-x-1}}{B(x, n-x)}$, $0 < \theta < 1$.

11.23. Resolver el problema 11.22 suponiendo la a priori de Jeffrey para la binomial, dada por $\pi(\theta) = \frac{1}{\sqrt{\theta(1-\theta)}}$, $0 < \theta < 1$.

$\pi(\theta|x) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \cdot \frac{1}{\theta^{1/2}(1-\theta)^{1/2}} = \binom{n}{x}\theta^{x-\frac{1}{2}}(1-\theta)^{n-x-\frac{1}{2}}$, $0 < \theta < 1$. Ésta es una densidad beta con $\alpha = x + \frac{1}{2}$ y $\beta = n - x + \frac{1}{2}$.

11.24. Suponer que se toma una muestra de una distribución exponencial (página 118) con parámetro θ desconocido que tiene la densidad a priori impropia $\pi(\theta) = 1/\theta$, $\theta > 0$. Determinar la densidad a posteriori $\pi(\theta|\boldsymbol{x})$.

$\pi(\theta|\boldsymbol{x}) \propto \theta^n e^{-\theta\sum_i x_i} \cdot \frac{1}{\theta} \propto \theta^{n-1}e^{-\theta\sum_i x_i}$, $\theta > 0$. En consecuencia, la densidad a posteriori para θ es gamma con parámetros $\alpha = n$ y $\beta = 1/\sum_i x_i$.

11.25. X es normal con media θ desconocida y varianza σ^2 conocida. La distribución a priori de θ es impropia y está dada por $\pi(\theta) = 1$, $-\infty < \theta < \infty$. Determinar la densidad a posteriori $\pi(\theta|\boldsymbol{x})$.

$\pi(\theta|\boldsymbol{x}) \propto e^{-\frac{1}{2\sigma^2}\sum_i (x_i-\theta)^2} \cdot 1 \propto e^{-\frac{n}{2\sigma^2}(\theta-\bar{x})^2}$. Por tanto, la distribución a posteriori es normal con media de $\bar{x}$ y varianza σ^2/n.

11.26. X es normal con media 0 y varianza θ no conocida. La varianza tiene la densidad a priori impropia $\pi(\theta) = 1/\sqrt{\theta}$, $\theta > 0$. Determinar la distribución a posteriori de θ.

$$\pi(\theta|x) \propto \frac{1}{\theta^{n/2}}e^{-\sum x^2/2\theta} \cdot \frac{1}{\sqrt{\theta}}, \theta > 0$$

$$\propto \theta^{-\left(\frac{n+1}{2}\right)}e^{-\sum x^2/2\theta}, \theta > 0$$

$$\propto \theta^{-\left(\frac{n-1}{2}\right)-1}e^{-\sum x^2/2\theta}, \theta > 0$$

Ésta es una densidad gamma inversa (vea el problema 11.99) con $\alpha = \dfrac{n-1}{2}$ y $\beta = \dfrac{\sum x^2}{2}$.

DISTRIBUCIONES A PRIORI CONJUGADAS

11.27. Una encuesta para predecir el resultado de un próximo referéndum determinó que 1 010 de 2 000 personas estuvieron a favor del referéndum. Suponiendo una densidad a priori uniforme de la proporción poblacional θ no conocida, determinar las probabilidades de que no se apruebe el referéndum. Comente el resultado con referencia a los problemas 11.11 y 11.12.

La distribución a posteriori de θ, dado el resultado de la encuesta, es beta con parámetros 1011, 991. Se necesita la probabilidad de $\theta < 0.5$, la cual resulta ser de 0.33, por lo que es posible tener una confianza de 33% de que el referéndum no se apruebe. Este resultado es el mismo que se obtuvo en el problema 11.12 utilizando para la distribución a priori la distribución a posteriori beta que obtuvimos en el problema 11.11. Puesto que la familia beta es conjugada con respecto a la distribución binomial, podemos actualizarla a posteriori de manera secuencial en el problema 11.12.

11.28. Una muestra aleatoria de tamaño 10 que se toma de una distribución geométrica con una probabilidad de éxito θ (vea la página 117) produce una media de 4.2. La densidad a priori de θ es uniforme en el intervalo [0, 1]. Determinar la distribución a posteriori de θ.

La distribución a priori es beta con parámetros $\alpha = \beta = 1$. Sabemos del ejemplo 11.12 que la distribución a posteriori también es beta. Los parámetros están dados por $\alpha + n = 11$ y $\beta + n\bar{x} - n = 33$.

11.29. Una muestra aleatoria de tamaño n se toma de una distribución binomial negativa con parámetro θ (vea la página 117): $f(x;\theta) = \dbinom{x-1}{r-1}\theta^r(1-\theta)^{x-r}$, $x = r, r+1, \ldots$. Se supone que la densidad a priori de θ es beta con parámetros α y β. Demostrar que la densidad a posteriori de θ es también beta, con parámetros $\alpha + nr$ y $\beta + n\bar{x} - nr$, donde $\bar{x}$ es la media muestral. Es decir, demostrar que la familia beta es conjugada con respecto a la distribución binomial negativa.

$\pi(\theta|\boldsymbol{x}) \propto \theta^{nr}(1-\theta)^{n(\bar{x}-r)} \cdot \theta^{\alpha-1}(1-\theta)^{\beta-1} \propto \theta^{nr+\alpha-1}(1-\theta)^{n(\bar{x}-r)+\beta-1}$, que es una densidad beta con parámetros $\alpha + nr$ y $\beta + n\bar{x} - nr$.

11.30. El tiempo entre llegadas de clientes en un banco está distribuido de manera exponencial con media de $1/\theta$, donde θ tiene una distribución gamma con parámetros $\alpha = 1$ y $\beta = 0.2$. Se observaron doce clientes durante un periodo y se determinó que el tiempo promedio entre llegadas fue de 6 minutos. Determinar la distribución a posteriori de θ.

Al aplicar el teorema 11.4, con $\alpha = 1$ y $\beta = 0.2$, $n = 11$ (12 clientes $\Rightarrow$ 11 intervalos entre llegadas), $\bar{x} = 6$, se ve que la densidad a posteriori es gamma con parámetros 12 y 0.014.

11.31. En el problema anterior, se supone que se observa una segunda muestra independiente, de 10 clientes y que se encuentra que tiene un tiempo promedio entre llegadas de 6.5 minutos. Determinar la distribución a posteriori de θ.

Puesto que la familia gamma es la conjugada de la distribución exponencial, es posible resolver este problema de dos formas: *i*) se empieza con la distribución a priori gamma con parámetros 1 y 0.2 y se aplica el teorema 11-4 con $n = 11 + 9 = 20$ y $\bar{x} = ((11 \times 6) + (9 \times 6.5))/20 \approx 6.225$ o bien *ii*) se comienza con la distribución a priori gamma con parámetros 12 y 0.014 y se aplica el teorema 11-4 con $n = 9$. Las dos formas llegan al resultado de que la densidad a posteriori es gamma con parámetros 21 y 0.0077.

11.32. La siguiente densidad se conoce como una *densidad de Rayleigh*: $f(x) = \theta x e^{-(\theta x^2/2)}$, $x > 0$. Es un caso especial de la densidad de Weibull (vea la página 118), con $b = 2$ y $a = \theta$. Demostrar que la familia gamma es conjugada con respecto a la distribución Rayleigh. Específicamente, demostrar que si X tiene una densidad de Rayleigh y θ una densidad a priori gamma con parámetros α y β, entonces la densidad a posteriori de θ dada una muestra aleatoria $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ de observaciones de X es también gamma.

$\pi(\theta|x) \propto f(x|\theta) \cdot \pi(\theta) \propto \theta^n e^{-\theta\sum_i x_i^2/2} \cdot \theta^{\alpha-1}e^{-\theta/\beta} \propto \theta^{(\alpha+n)-1}e^{-\theta\left(\frac{1}{\beta}+\sum_i x_i^2/2\right)}$, $\theta > 0$. Ésta es una densidad gamma con parámetros $\alpha + n$ y $\dfrac{2\beta}{2 + \beta\sum x^2}$.

11.33. Demostrar que la familia gamma inversa (vea el problema 11.99) es conjugada con respecto a la distribución normal con media conocida pero varianza θ desconocida.

Se supone que la media de la densidad normal es 0. Tenemos

$$f(\boldsymbol{x}\,|\,\theta) = \frac{1}{(2\pi)^{n/2}\theta^{n/2}}\exp\left\{-\frac{1}{2\theta}\sum_{i=1}^{n}x_i^2\right\} \qquad \text{y}$$

$$\pi(\theta) = \frac{\beta^{\alpha}\theta^{-\alpha-1}e^{-\beta/\theta}}{\Gamma(\alpha)},\ \theta > 0$$

La densidad a posteriori está dada por

$$\pi(\theta\,|\,\boldsymbol{x}) = f(\boldsymbol{x}\,|\,\theta)\cdot\pi(\theta) \propto \theta^{-n/2}e^{-\frac{1}{2\theta}\sum_i x_i^2}\cdot\theta^{-\alpha-1}e^{-\beta/\theta} \propto \theta^{(-\frac{n}{2}-\alpha)-1}e^{-\frac{1}{\theta}\left(\beta+\frac{\sum_i x_i^2}{2}\right)},\ \theta > 0$$

Ésta también es una densidad gamma inversa, con parámetros $\dfrac{n}{2}+\alpha$ y $\beta + \dfrac{\sum_i x_i^2}{2}$.

11.34. Se toma una muestra aleatoria de n observaciones de la densidad exponencial con media θ: $f(\boldsymbol{x}\,|\,\theta) = (1/\theta)\exp\{-x/\theta\},\ x > 0$. Suponer que θ tiene una distribución a priori gamma inversa (vea el problema 11.99) y demostrar que su distribución a posteriori también es de la familia gamma inversa.

$$f(\boldsymbol{x}\,|\,\theta) = (1/\theta)^n\exp\left\{-\sum_i x_i/\theta\right\},\ x > 0$$

$$\pi(\theta) = \frac{\beta^{\alpha}\theta^{-\alpha-1}e^{-\beta/\theta}}{\Gamma(\alpha)},\ \theta > 0$$

La densidad a posteriori, dada por $\pi(\theta\,|\,\boldsymbol{x}) \propto f(\boldsymbol{x}\,|\,\theta)\cdot\pi(\theta) \propto \theta^{-n}e^{-\frac{\sum_i x_i}{\theta}}\cdot\theta^{-\alpha-1}e^{-\beta/\theta} \propto \theta^{-(n+\alpha)-1}e^{-\frac{1}{\theta}\left(\beta+\sum x_i\right)},\ \theta > 0$, es una densidad gamma inversa con parámetros $n+\alpha$ y $\beta + \sum_i x_i$.

11.35. En el problema anterior, se supone que una segunda muestra de m observaciones de la misma población produce las observaciones $y_1, y_2, \ldots, y_m$. Determinar la densidad a posteriori previa incorporación del resultado de las dos muestras.

En razón de que la familia gamma inversa es conjugada con respecto a la distribución exponencial, se pueden actualizar los parámetros que se obtuvieron en el problema 11.34 como $m + (n+\alpha)$ y $\left(\beta + \sum_i x_i\right) + \sum_j y_j$.

En consecuencia, la densidad a posteriori es la gamma inversa con parámetros $m+n+\alpha$ y $\beta + \sum_i x_i + \sum_j y_j$.

11.36. Se toma una muestra aleatoria de n observaciones de la densidad gamma:

$f(x\,|\,\theta) = \dfrac{x^{\alpha-1}e^{-x/\theta}}{\theta^{\alpha}\Gamma(\alpha)},\ x > 0$. Suponer que θ tiene una distribución a priori gamma inversa con parámetros γ y β, y demostrar que su distribución a posteriori también es de la familia gamma inversa.

$$f(\boldsymbol{x}\,|\,\theta) \propto (1/\theta^{\alpha})^n\exp\left\{-\sum_i x_i/\theta\right\},\ x > 0$$

$$\pi(\theta) = \frac{\beta^{\gamma}\theta^{-\gamma-1}e^{-\beta/\theta}}{\Gamma(\gamma)},\ \theta > 0$$

La densidad a posteriori, dada por $\pi(\theta\,|\,x) \propto f(\boldsymbol{x}\,|\,\theta)\cdot\pi(\theta) \propto \theta^{-n\alpha}e^{-\frac{-\sum_i x_i}{\theta}}\cdot\theta^{-\gamma-1}e^{-\beta/\theta} \propto \theta^{-(n\alpha+\gamma)-1}e^{-\frac{1}{\theta}\left(\beta+\sum_i x_i\right)},\ \theta > 0$ es gamma inversa con parámetros $n\alpha+\gamma$ y $\beta + \sum_i x_i$.

ESTIMACIÓN PUNTUAL BAYESIANA

11.37. En el problema 11.5, calcular la estimación de Bayes de ρ con la función de pérdida del error cuadrático.

a) La estimación de Bayes es la media de la distribución a posteriori, que corresponde a

$$\sum_{\rho=1}^{n}\rho\cdot\binom{n-1}{\rho-1}p^{\rho-1}(1-p)^{n-\rho} = 1 + \sum_{\rho=1}^{n}(\rho-1)\cdot\binom{n-1}{\rho-1}p^{\rho-1}(1-p)^{n-\rho} = 1 + (n-1)p$$

b) La media a posteriori es $\sum_{\rho=0}^{n-1} \rho \cdot \binom{n-1}{\rho} p^{\rho}(1-p)^{n-1-\rho} = (n-1)p$.

11.38. Demostrar que la estimación de Bayes μ_{post} para θ que se obtuvo en el teorema 11-5 es una combinación convexa del estimador de máxima verosimilitud de θ y la media a priori de θ.

$$\mu_{\text{post}} = \frac{x+\alpha}{n+\alpha+\beta} = \left(\frac{n}{n+\alpha+\beta}\right)\left(\frac{x}{n}\right) + \left(\frac{\alpha+\beta}{n+\alpha+\beta}\right)\left(\frac{\alpha}{\alpha+\beta}\right)$$

11.39. En el problema 11.10, determinar la estimación de Bayes con la función de pérdida del error cuadrático para *a*) θ *b*) $1/\theta$.

a) La distribución a posteriori es beta con parámetros 25 y 17. La estimación de Bayes, que equivale a la media a posteriori, es igual a $25/52 \approx 0.48$.

b) La estimación de Bayes de $1/\theta$ es la media a posteriori de $1/\theta$, dada por

$$\frac{1}{B(25,17)}\int_0^1 \frac{1}{\theta} \cdot \theta^{24}(1-\theta)^{16} d\theta = \frac{B(24,17)}{B(25,17)} = \frac{41}{24} \approx 1.71$$

11.40. En el problema 11.15, calcular la estimación de Bayes con la función de pérdida del error cuadrático para θ.

La estimación de Bayes es la media a posteriori, la cual de acuerdo con el problema 11.15 se sabe que es 4.38.

11.41. En el problema 11.33, suponer que $\alpha = \beta = 1$ y determinar la estimación de Bayes para la varianza con la pérdida del error cuadrático.

La distribución a posteriori es gamma inversa (vea el problema 11.99) con parámetros $\frac{n}{2} + 1$ y $1 + \frac{\sum_i x_i^2}{2}$. La estimación de Bayes es la media a posteriori, dada por $\frac{2 + \sum_i x_i^2}{n}$.

11.42. Determinar la estimación de Bayes de θ con la función de pérdida del error cuadrático del problema 11.24 y compararla con el estimador de máxima verosimilitud.

Los parámetros de la a posteriori son n y $1 \Big/ \sum_i x_i$. Por tanto, la estimación de Bayes, que es la media a posteriori, es $1/\bar{x}$ lo que es igual al estimador de máxima verosimilitud para θ (vea el problema 11.98).

11.43. En el ejemplo 11.10, determinar la estimación de Bayes para θ bajo la función de pérdida del error cuadrático.

La distribución a posteriori de θ es normal con media $\bar{x}$ y varianza σ^2/n. La estimación de Bayes de θ bajo la pérdida del error cuadrático, que es la media a posteriori, está dada por $\bar{x}$.

11.44. En el problema 11.30, determinar la estimación de Bayes para θ con la función de pérdida del error cuadrático. Calcular la pérdida del error cuadrático del estimador para cada $x = (x_1, x_2, \ldots, x_n)$ y compararla con la pérdida del estimador de máxima verosimilitud.

La estimación de Bayes bajo la pérdida del error cuadrático es la media a posteriori $= \beta(\alpha + n)/(1 + n\beta\bar{x})$. La pérdida del error cuadrático para cada $\boldsymbol{x}$ es la varianza a posteriori $\beta^2(\alpha + n)/(1 + n\beta\bar{x})^2$. Con $\alpha = 1$ y $\beta = 0.2$, $n = 11$ y $\bar{x} = 6$, el resultado es 0.00238. El estimador de máxima verosimilitud para θ es $1/\bar{x}$ y su pérdida del error cuadrático corresponde a $E\left[\left(\frac{1}{\bar{x}} - \theta\right)^2 \middle| \boldsymbol{x}\right] = \left(\frac{1}{\bar{x}}\right)^2 - \frac{2}{\bar{x}} E(\theta|x) + E(\theta^2|x)$. Con $\alpha = 1$ y $\beta = 0.2$, $n = 11$ y $\bar{x} = 6$, el resultado de la expresión anterior es 0.00239.

11.45. Si X es una variable aleatoria de Poisson con parámetro λ, cuya densidad a priori de λ es gamma con parámetros α y β, demostrar que la estimación de Bayes para λ es un promedio ponderado de su estimador de máxima verosimilitud y la media a priori.

Según el teorema 11-2, la distribución a posteriori es gamma con parámetros $n\bar{x} + \alpha$ y $\beta/(1 + n\beta)$. La media a posteriori es $\frac{\beta(n\bar{x} + \alpha)}{(1 + n\beta)} = \frac{n\beta}{1 + n\beta} \cdot \bar{x} + \frac{1}{1 + n\beta} \cdot \alpha\beta$.

11.46. En el problema 11.16, determinar la estimación de Bayes de θ con *a*) la pérdida del error cuadrático y *b*) la pérdida del error absoluto.

a) La estimación de Bayes con la pérdida del error cuadrático es la media a posteriori de θ, que es 1.17.
b) La estimación de Bayes con la pérdida del error absoluto es la mediana a posteriori, la cual es igual a la media a posteriori, en este caso, puesto que la distribución a posteriori es normal.

11.47. En el problema 11.32, determinar la estimación de Bayes para θ con la función de pérdida del error cuadrático.

La distribución a posteriori de θ es gamma con parámetros $\alpha + n$ y $\dfrac{2\beta}{2 + \beta\sum x^2}$. Por tanto, la media a posteriori es $\dfrac{2\beta(\alpha + n)}{\left(2 + \beta\sum_i x_i^2\right)}$.

11.48. El tiempo (en minutos) que un cliente de un banco tiene que esperar en una cola para que lo atiendan está distribuido exponencialmente con media $1/\theta$. La distribución a priori de θ es gamma con media de 0.4 y desviación estándar de 1. En una muestra aleatoria de 10 clientes se registraron los siguientes tiempos de espera: 2, 3.5, 1, 5, 4.5, 3, 2.5, 1, 1.5, 1. Calcular la estimación de Bayes para θ con la función de pérdida *a*) del error cuadrático y *b*) del error absoluto.

La distribución gamma con parámetros α y β tiene media $\alpha\beta$ y varianza $\alpha\beta^2$. En consecuencia, los parámetros para esta gamma a priori deben ser $\alpha = 0.16$ y $\beta = 2.5$. La distribución a posteriori es (vea el teorema 11-4) gamma con parámetros $\alpha + n = 10.16$ y $\beta/(1 + n\beta\bar{x}) = 0.04$.

a) La media a posteriori corresponde a $10.16 \times 0.04 = 0.41$.
b) La mediana de la densidad a posteriori, que se obtuvo mediante software de computadora, es igual a 0.393.

11.49. En el problema 11.6, determinar la estimación de Bayes con la pérdida del error cuadrático para θ en cada caso y evaluarlo si $n = 500$ y $x = 200$.

a) A partir del teorema 11-6 se sabe que, en este caso, la estimación de Bayes es la media a posteriori. La media de la densidad beta con parámetros $x + 2$ y $n - x + 1$ es $(x + 2)/(n + 3) = 0.4016$.
b) De manera similar a lo anterior, la estimación de Bayes es $(x + 3)/(n + 4) = 0.4028$.
c) La estimación de Bayes es $(x + 4)/(n + 5) = 0.4040$.

11.50. En el problema 11.6, parte *a*), calcular la estimación de Bayes con la pérdida del error cuadrático para la desviación estándar poblacional, $\sqrt{n\theta(1 - \theta)}$.

El estimador que se busca es la esperanza a posteriori de $\sqrt{n\theta(1 - \theta)}$, la cual es igual a

$$\frac{\sqrt{n}}{B(x + 2, n - x + 1)}\int_0^1 \theta^{\frac{1}{2}}(1 - \theta)^{\frac{1}{2}} \cdot \theta^{x+1}(1 - \theta)^{n-x} d\theta = \sqrt{n} \times \frac{B\left(x + \frac{5}{2}, n - x + \frac{3}{2}\right)}{B(x + 2, n - x + 1)}$$

11.51. En el problema 11.6, determinar la estimación de Bayes con la pérdida del error absoluto para θ en cada caso bajo el supuesto de que $n = 500$ y $x = 200$.

Según el teorema 11-6, la estimación de Bayes de θ con la pérdida del error absoluto es la mediana de la distribución a posteriori de θ. Puesto que no existe una expresión explícita para la mediana de una densidad beta, se obtuvieron los siguientes valores de medianas con un software de computadora: *a*) 0.4015; *b*) 0.4027; *c*) 0.4038.

11.52. En el problema 11.14, estimar λ utilizando la estimación de Bayes con *a*) la pérdida del error cuadrático y *b*) la pérdida del error absoluto.

La densidad a posteriori que se obtuvo en el problema 11.14 es una densidad gamma con parámetros 25 y 0.091.

a) La estimación de Bayes con la pérdida del error cuadrático es la media a posteriori que, en este caso, corresponde a $\alpha\beta = 2.275$.
b) La estimación de Bayes con la pérdida del error absoluto es la mediana a posteriori. Mediante software de computadora para calcular la mediana de la distribución a posteriori gamma se obtiene la estimación 2.245.

11.53. Una muestra aleatoria $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ de tamaño n se toma de una población con la función de densidad $f(x|\theta) = 3\theta x^2 e^{-\theta x^3}$, $0 < x < \infty$, donde θ tiene una densidad a priori gamma con parámetros α y β. Determinar la estimación de Bayes para θ con la pérdida del error cuadrático.

$\pi(\theta|\boldsymbol{x}) \propto \theta^n e^{-\theta\sum x^3} \cdot \theta^{\alpha-1}e^{-\theta/\beta} \propto \theta^{(n+\alpha)-1}e^{-\theta\left(\sum x^3+\frac{1}{\beta}\right)}$, que es una densidad gamma con parámetros $\alpha + n$ y $\dfrac{\beta}{1+\beta\sum x^3}$.

En consecuencia, la estimación de la media a posteriori de θ es $\dfrac{\beta(\alpha+n)}{1+\beta\sum x^3}$.

11.54. En el problema 11.24, calcular la estimación de Bayes de $e^{-t\theta}$ con respecto a la función de pérdida del error cuadrático.

La estimación de Bayes es $E(e^{-t\theta}|\boldsymbol{x}) = \displaystyle\int_0^\infty e^{-t\theta}\pi(\theta|\boldsymbol{x})\,d\theta = \int_0^\infty e^{-t\theta}\theta^{n-1}e^{-\theta\sum_i x_i}d\theta = \int_0^\infty \theta^{n-1}e^{-\theta(t+\sum_i x_i)}d\theta = \dfrac{\Gamma(n)}{(t+n\bar{x})^n}$.

11.55. La variable aleatoria X está distribuida de manera normal con media θ y varianza σ^2. La distribución a priori de θ es normal estándar. *a*) Determinar la estimación de Bayes de θ con la función de pérdida del error cuadrático basada en una muestra aleatoria de tamaño n. *b*) ¿El estimador que resulta es insesgado (vea la página 195)? *c*) Comparar la estimación de Bayes con la estimación de máxima verosimilitud en términos de la pérdida del error cuadrático.

a) Según el teorema 11-3, la estimación de Bayes es $\left(\dfrac{\bar{\ }}{+\sigma^2}\right)$. Con $c = \left(\dfrac{\ }{+\sigma^2}\right)$, la pérdida del error cuadrático de la estimación anterior es $E[(c\bar{x}-\theta)^2|\boldsymbol{x}] = c^2\bar{x}^2 - 2c\bar{x}\cdot 0 + 1 = c^2\bar{x}^2 + 1$.

b) Puesto que $E\left(\dfrac{n\bar{X}}{n+\sigma^2}\right) = \dfrac{n\theta}{n+\sigma^2}$, el estimador es sesgado. Sin embargo, es insesgado asintóticamente.

c) La estimación de máxima verosimilitud de θ es $\bar{x}$. La pérdida del error cuadrático para esta estimación es $\bar{x}^2 + 1$. Es obvio, puesto que $c < 1$, la pérdida es menor que la estimación de Bayes. Para valores grandes de n, las pérdidas son aproximadamente iguales.

11.56. En el problema 11.22, demostrar que la estimación de Bayes de θ es igual a la estimación de máxima verosimilitud.

La estimación de Bayes es la media a posteriori de θ, dada por $\dfrac{\alpha}{\alpha+\beta} = \dfrac{x}{n}$. La estimación de máxima verosimilitud se encuentra cuando se maximiza la verosimilitud $L \propto \theta^x(1-\theta)^{n-x}$ con respecto a θ (vea la página 198). Si resolvemos la ecuación $\dfrac{dL}{d\theta} = x\theta^{x-1}(1-\theta)^{n-x} + (n-x)\theta^x(1-\theta)^{n-x-1} = 0$ para θ, obtenemos el estimador de máxima verosimilitud x/n.

11.57. En el problema 11.48, determinar la estimación de Bayes para $1/\theta$ con la función de pérdida del error cuadrático.

La estimación de Bayes es la esperanza de $1/\theta$ con respecto a la distribución a posteriori de θ:

$$E\left(\frac{1}{\theta}|\boldsymbol{x}\right) = \frac{1}{(0.04)^{10.16}\Gamma(10.16)}\int_0^\infty \frac{1}{\theta}\cdot\theta^{9.16}e^{-\theta/0.04}d\theta = \frac{(0.04)^{9.16}\Gamma(9.16)}{(0.04)^{10.16}\Gamma(10.16)} \approx 2.73$$

ESTIMACIÓN POR INTERVALO BAYESIANA

11.58. Las mediciones de los diámetros de una muestra aleatoria de 200 cojinetes de bolas fabricados por una máquina particular, durante una semana, indicaron una media de 0.824 pulgadas y una desviación estándar de 0.042 pulgadas. Los diámetros se distribuyen normalmente. Calcular el intervalo de credibilidad bayesiano de MDP a *a*) 90%, *b*) 95% y *c*) 98% para el diámetro medio θ de todos los cojinetes de bolas fabricados por la máquina. Se supone que la distribución a priori de θ es normal con media de 0.8 pulgadas y desviación estándar de 0.05.

La media a posteriori y la desviación estándar son 0.824 y 0.0030, respectivamente.

a) El intervalo de MDP de 90% está dado por $0.824 \pm (1.645 \times 0.003)$ o bien [0.819, 0.829].
b) El intervalo de MDP de 95% está dado por $0.824 \pm (1.96 \times 0.003)$ o bien [0.818, 0.830].
c) El intervalo de MDP de 98% corresponde a $0.824 \pm (2.33 \times 0.003)$ o bien [0.817, 0.831].

11.59. Una encuesta muestral de 100 votantes elegidos al azar de todo el padrón de un distrito indicó que 55% de ellos estuvieron a favor de cierto candidato. Se supone que, previo a la encuesta, se cree que la proporción

verdadera θ de votantes en ese distrito a favor del candidato tiene una a priori tipo Jeffrey (vea el problema 11.23) dada por $\pi(\theta) = \dfrac{1}{\sqrt{\theta(1-\theta)}}$, $0 < \theta < 1$. Determinar intervalos de credibilidad bayesianos del área de colas iguales a 95 y 99% para la proporción θ de todos los votantes a favor de dicho candidato.

Se tiene $n = 100$ y $x = 55$. A partir del problema 11.23, la densidad a posteriori de θ es beta con parámetros de $\alpha = 55.5$ y $\beta = 45.5$. Esta densidad tiene los siguientes percentiles: $x_{0.005} = 0.423$, $x_{0.025} = 0.452$, $x_{0.975} = 0.645$, $x_{0.995} = 0.673$. Ello origina el intervalo de credibilidad bayesiano [0.452, 0.645] de colas iguales a 95% y el intervalo de credibilidad bayesiano [0.423, 0.673] de colas iguales a 99%. (Es instructivo comparar los intervalos anteriores con los tradicionales que se obtuvieron en el problema 6.13.)

11.60. En el problema anterior, asumir que θ tiene una distribución a priori uniforme sobre [0,1] y calcular los intervalos de credibilidad del área de colas iguales a *a*) 95% y *b*) 99% para θ.

La distribución a posteriori de θ es beta con parámetros 56 y 46 (vea el teorema 11-1).

a) Se necesitan los percentiles $x_{0.025}$ y $x_{0.975}$ de la distribución beta anterior, que son 0.452 y 0.644, respectivamente. El intervalo a 95% es [0.452, 0.644].

b) Se necesitan los percentiles $x_{0.005}$ y $x_{0.995}$ de la distribución beta anterior, que son 0.422 y 0.644, respectivamente. El intervalo a 99% es [0.422, 0.672].

11.61. En 40 lanzamientos de una moneda se obtuvieron 24 caras. Determinar el intervalo de credibilidad de 90 y 99.73% para la proporción de caras θ que se obtendría en un número ilimitado de lanzamientos de la moneda. Utilice una a priori uniforme para θ.

Según el teorema 11-1, la densidad a posteriori de θ es beta con $\alpha = 25$ y $\beta = 17$. Esta densidad tiene los siguientes percentiles de $x_{0.00135} = 0.367$, $x_{0.05} = 0.469$, $x_{0.95} = 0.716$, $x_{0.99865} = 0.800$. Los intervalos de credibilidad bayesianos del área de dos colas iguales a 90 y 99.73% son, respectivamente, [0.469, 0.716] y [0.367, 0.800]. (Los intervalos de confianza tradicionales están dados en el problema 6.15.)

11.62. En una muestra de 100 mediciones del diámetro de una esfera se obtuvo una media de $\bar{x} = 4.38$ pulgadas. Con base en una experiencia anterior, se sabe que el diámetro está distribuido de manera normal con media θ desconocida y varianza 0.36. *a*) Calcular los intervalos de credibilidad del área de colas iguales a 95 y 90% para el diámetro θ verdadero bajo el supuesto de una densidad a priori normal con media de 4.5 pulgadas y varianza de 0.4. *b*) ¿Con qué grado de credibilidad se puede decir que el diámetro verdadero es 4.38 ± 0.01?

a) Del teorema 11-3 se observa que la media y la varianza ambas a posteriori para θ son 4.381 y 0.004, respectivamente. El intervalo de credibilidad a 95% es $[4.381 - (1.96 \times 0.063), 4.381 + (1.96 \times 0.063] = [4.26, 4.50]$. De manera similar, el intervalo de credibilidad a 90% corresponde a $[4.381 - (1.645 \times 0.063), 4.381 + (1.645 \times 0.063] = [4.28, 4.48]$.

b) Se necesita el área bajo la densidad a posteriori de 4.37 a 4.39, que es igual al área bajo la densidad normal estándar entre $(4.37 - 4.381)/0.063 = -0.17$ y $(4.39 - 4.381)/0.063 = 0.14$, lo que es igual a 0.1232, así que el grado de credibilidad obtenido es de, aproximadamente, 12%.

11.63. En el problema 11.16, construir un intervalo de credibilidad a 95% para θ.

De acuerdo con el problema 11.16, vemos que la media y la varianza ambas a posteriori para θ son 1.17 y 0.006, respectivamente. El intervalo de credibilidad a 95% es $[1.17 - (1.96 \times 0.077), 1.17 + (1.96 \times 0.077] = [1.02, 1.32]$.

11.64. En el problema 11.25, ¿qué se puede decir acerca del intervalo de credibilidad bayesiano *MDP* para θ en comparación con el intervalo convencional que se muestra en (1) en el capítulo 6?

La distribución a posteriori de θ es normal con media $\bar{x}$ y varianza σ^2/n. Los intervalos de credibilidad *MDP* que se obtienen serían idénticos a los intervalos de confianza convencionales centrados en $\bar{x}$.

11.65. El número de individuos que anualmente padecen una reacción adversa por una inyección de suero tiene una distribución de Poisson con media λ desconocida. Se supone que λ tiene una densidad a priori impropia tipo Jeffrey $\pi(\lambda) = 1/\sqrt{\lambda}$, $\lambda > 0$ (vea el problema 11.21). La tabla 11-8 proporciona el número de este tipo de casos que ocurrieron en cada uno de los últimos 10 años.

a) Determinar el intervalo de credibilidad de colas iguales a 98% para λ. *b*) ¿Con qué grado de credibilidad es posible afirmar que λ no excede de 3?

Tabla 11-8

Año	1	2	3	4	5	6	7	8	9	10
Número	2	4	1	2	2	1	2	3	3	0

a) Del problema 11.21 sabemos que la distribución a posteriori para λ es gamma con parámetros $n\bar{x} + \frac{1}{2}$ y $\frac{1}{n}$, que en este caso son 20.5 y 0.1. En consecuencia, se necesitan el 1º y 99º percentil de la distribución gamma con estos parámetros. Mediante software de computadora, obtenemos $x_{0.01} = 1.146$ y $x_{0.99} = 3.248$. El intervalo de credibilidad al 98% es [1.146, 3.248].

b) Se necesita que la probabilidad a posteriori de λ no exceda de 3. Ésta corresponde al área a la izquierda de 3 bajo la densidad gamma con parámetros 20.5 y 0.1. Puesto que esta área es 0.972, es posible tener la confianza de 97% de que λ no exceda de 3.

11.66. En el problema 11.14, obtener el intervalo de credibilidad bayesiano del área de colas iguales a 95% para λ.

La densidad a posteriori se obtuvo en el problema 11.14 como una densidad gamma con parámetros 25 y 0.091. Los percentiles de esta densidad relevantes para el intervalo de credibilidad de este caso son $x_{0.975} = 3.25$ y $x_{0.025} = 1.47$. El intervalo de credibilidad bayesiano a 95% es [1.47, 3.25].

11.67. Obtener un intervalo de credibilidad de colas iguales a 95% para θ en el problema 11.22 bajo el supuesto de que $n = 10$, $x = 3$.

La a posteriori es beta con parámetros 3 y 7. Los percentiles son $x_{0.025} = 0.075$ y $x_{0.975} = 0.600$. El intervalo es [0.075, 0.600].

11.68. Determinar el intervalo de credibilidad de colas iguales a 95% para θ del problema 11.23 suponiendo que $n = 10$, $x = 3$.

La a posteriori es beta con parámetros 3.5 y 7.5. Los percentiles son $x_{0.025} = 0.093$ y $x_{0.975} = 0.606$. El intervalo es [0.093, 0.606].

11.69. En el problema 11.48, obtener el intervalo de credibilidad del área de colas iguales a 99% para *a*) θ y *b*) $1/\theta$.

a) La distribución a posteriori de θ es gamma con parámetros 10.16 y 0.04. Se obtuvieron los siguientes percentiles de la distribución anterior mediante el uso de software de computadora: $x_{0.005} = 0.15$ y $x_{0.995} = 0.81$. El intervalo de credibilidad es [0.15, 0.81].

b) Puesto que $\theta < 0.15 \Leftrightarrow 1/\theta > 1/0.15$ y $\theta > 0.81 \Leftrightarrow 1/\theta < 1/0.81$, el intervalo del área de colas iguales para $1/\theta$ es $[1/0.81, 1/0.15] = [1.23, 6.67]$.

PRUEBAS DE HIPÓTESIS BAYESIANAS

11.70. Se sabe que el tiempo de vida media (en horas) de focos fluorescentes que produce una compañía corresponde a una distribución normal con una media θ no conocida y una desviación estándar de 120 horas. La densidad a priori de θ es normal con $\mu = 1\,580$ horas y $v^2 = 16\,900$. Un tiempo de vida media de una muestra de 100 focos se calcula igual a 1 630 horas. Probar la hipótesis nula H_0: $\theta \leq 1\,600$ contra la hipótesis alternativa H_1: $\theta > 1\,600$ usando pruebas de Bayes de *a*) 0.05 y *b*) 0.01.

a) Según el teorema 11.3, la densidad a posteriori es normal con una media de 1629.58 y una desviación estándar de 11.95. La probabilidad a posteriori de H_0 es $P(\theta \leq 1600|\boldsymbol{x}) = P\left(Z \leq \dfrac{1600 - 1629.58}{11.95}\right) \approx 0.007$ y puesto que es menor que 0.05, se rechaza H_0.

b) Como la probabilidad a posteriori de la hipótesis nula, que se obtuvo en *a*), es menor que 0.01, se rechaza H_0.

11.71. Si en el ejemplo 11.18 una segunda muestra de 100 observaciones produjo un tiempo de reacción media de 0.25 seg. Probar la hipótesis nula H_0: $\theta \leq 0.3$ contra la alternativa H_1: $\theta > 0.3$ utilizando una prueba de Bayes de 0.05.

Se considera que la distribución a priori de θ es la distribución a posteriori que se obtuvo en el ejemplo 11.18: normal con media 0.352 y varianza 0.004. Con base en el teorema 11-3 con esta a priori y los datos nuevos, se obtienen una media y varianza ambas a posteriori de 0.269 y 0.0007. La probabilidad a posteriori de la hipótesis nula es 0.88, y puesto que no es menor que 0.05, no es posible rechazar la hipótesis nula.

11.72. En el problema 11.21 se supone que una muestra de tamaño 10 produce los valores 2, 0, 1, 1, 3, 0, 2, 4, 2, 2. Probar $H_0 : \lambda \leq 0.1$ contra $H_1 : \lambda > 1$ mediante una prueba de Bayes de 0.05.

Se necesita la probabilidad a posteriori de H_0, la cual, según el problema 11.21, es el área de 0 a 1 bajo la densidad gamma con parámetros $n\bar{x} + \frac{1}{2} = 17.5$ y $\frac{1}{n} = 0.1$. Con software de computadora, observamos que esta probabilidad es 0.02. En vista de que dicho valor es menor del umbral especificado de 0.05, rechazamos la hipótesis nula.

11.73. En el problema 11.65 probar la hipótesis nula $H_0 : \lambda \leq 0.1$ contra la alternativa $H_1 : \lambda > 1$ usando una prueba de Bayes de 0.05.

La prueba de Bayes de 0.05 rechazaría la hipótesis nula si la probabilidad a posteriori de la hipótesis $\lambda \leq 1$ es menor que 0.05. En este caso, la probabilidad anterior está dada por el área a la izquierda de 1 bajo la distribución gamma con parámetros 20.5 y 0.1 y es 0.002. Puesto que esto es menor que 0.05 se rechaza la hipótesis nula.

11.74. En el problema 11.6 se supone que $n = 40$ y $x = 10$, probar la hipótesis nula $H_0 : \theta \leq 0.2$ contra la alternativa $H_1 : \theta > 0.2$ utilizando una prueba de Bayes de 0.05.

a) La probabilidad a posteriori de la hipótesis nula está dada por el área de 0 a 0.2 bajo una densidad beta con parámetros 12 y 31, la que se determina es igual a 0.12 por medio de software de computadora. Puesto que lo anterior no es menor que 0.05, no se puede rechazar la hipótesis nula.

b) La probabilidad a posteriori es el área de 0 a 0.2 bajo la densidad beta con parámetros 13 y 31, que es 0.07, y, puesto que no es menor que 0.05, no se puede rechazar la hipótesis nula.

c) La probabilidad a posteriori es el área de 0 a 0.2 bajo una densidad beta con parámetros 14 y 31, que es 0.04. Como este valor es menor que 0.05 se rechaza la hipótesis nula.

11.75. En el problema 11.48 probar la hipótesis nula $H_0 : \theta \geq 0.7$ contra $H_1 : \theta < 0.7$ mediante una prueba de Bayes de 0.025.

La distribución a posteriori de θ es gamma con parámetros 10.16 y 0.04. Por tanto, la probabilidad a posteriori de la hipótesis nula es 0.022. Debido a que ello es menor que 0.025, se rechaza la hipótesis nula.

11.76. El tiempo de vida X de un componente de cómputo tiene la densidad exponencial dada por (vea la página 118) $f(x \mid \theta) = \theta e^{-\theta x}$, $x > 0$ con media $1/\theta$ desconocida. Se supone que la densidad a priori de θ es gamma con parámetros $\alpha = 0.2$ y $\beta = 0.15$. Si una muestra aleatoria de 10 observaciones en X produjo un tiempo de vida promedio de 7 años, utilizar la prueba de Bayes de 0.05 para probar la hipótesis nula de que el tiempo de vida esperada es, por lo menos, de 12 años contra la hipótesis alternativa de que es menor que 12 años.

Las hipótesis nula y alternativa corresponden, respectivamente, a $H_0 : \theta \leq 1/12 = 0.083$ y $H_1 : \theta > 0.083$. A partir del teorema 11-4, la distribución a posteriori de θ es gamma con parámetros 10.2 y 0.013. La probabilidad a posteriori de la hipótesis nula es 0.10. Debido a que el resultado anterior es mayor que 0.05, no se puede rechazar la hipótesis nula.

FACTOR DE BAYES

11.77. En el ejemplo 11.4, determinar el factor de Bayes de $H_0 : \lambda = 1$ relativo a $H_1 : \lambda \neq 1$.

$$FB = \{P(H_0|x)/[1 - P(H_0|x)]\} \div \{P(H_0)/[1 - P(H_0)]\} = (0.49/0.51) \div ((1/3)/(2/3)) \approx 1.92$$

11.78. Es deseable probar la hipótesis nula $\theta \leq 0.6$ contra la alternativa $\theta > 0.6$, donde θ es la probabilidad de éxito para un ensayo de Bernoulli. Se supone que θ tiene una distribución a priori uniforme en [0,1] y que en 40 ensayos ocurrieron 24 éxitos. ¿Cuál es la conclusión si se decide rechazar la hipótesis nula si $FB < 1$?

La densidad a posteriori de θ es beta con $\alpha = 25$ y $\beta = 17$. La probabilidad a posteriori de la hipótesis nula es 0.52. La razón de posibilidades a posteriori es $0.52/0.48 = 1.0833$ y el a priori $6/4 = 1.5$, $FB = 0.72$. Se rechaza la hipótesis nula.

11.79. Demostrar que la regla *ad hoc* (vea el comentario después del teorema 11-10) para rechazar H_0 si $FB \leq 1$, es equivalente a la prueba α de Bayes con $\alpha = P(H_0)$.

$$FB \leq 1 \Leftrightarrow \left(\frac{P(H_0|\boldsymbol{x})}{P(H_1|\boldsymbol{x})}\right) \Big/ \left(\frac{P(H_0)}{P(H_1)}\right) \leq 1 \Leftrightarrow P(H_0|\boldsymbol{x})[1 - P(H_0)] \leq [1 - P(H_0|\boldsymbol{x})]P(H_0) \Leftrightarrow P(H_0|\boldsymbol{x}) \leq P(H_0)$$

11.80. En el problema anterior calcular c de modo que el criterio del factor de Bayes para rechazar la hipótesis nula si $FB < c$ equivalga a la regla de Bayes de 0.05.

Según el teorema 11-10, $c = \dfrac{\alpha[1 - P(H_0)]}{(1 - \alpha)P(H_0)} = \dfrac{(0.05)(1 - 0.6)}{(1 - 0.05)(0.6)} \approx 0.035$.

11.81. Resolver el problema 11.71 con base en la decisión de rechazar la hipótesis nula si el factor de Bayes es menor que 1. Por el problema 11.79 se sabe que la regla para rechazar H_0 si $FB < 1$ es equivalente a rechazar la hipótesis nula si $P(H_0|\boldsymbol{x}) \leq P(H_0)$. De acuerdo con el problema 11.71, $P(H_0|\boldsymbol{x}) = 0.88$. Del ejemplo 11.18, se sabe que la distribución a priori de θ es normal con media 0.4 y varianza 0.13. En consecuencia, $P(H_0) = P\left(Z \leq \dfrac{0.3 - 0.4}{0.361}\right) \approx 0.39$, por lo que no podemos rechazar la hipótesis nula.

11.82. En el problema 11.74, probar cada uno de los casos utilizando la regla del factor de Bayes para rechazar la hipótesis nula si $FB \leq 4$.

a) $FB = \{P(H_0|x)/[1 - P(H_0|x)]\} \div \{P(H_0)/[1 - P(H_0)]\} = (0.12/0.88) \div (0.04/0.96) \approx 3.27$. Se rechaza la hipótesis nula.

b) $FB = \{P(H_0|x)/[1 - P(H_0|x)]\} \div \{P(H_0)/[1 - P(H_0)]\} = (0.07/0.93) \div (0.008/0.992) \approx 9.33$. No se puede rechazar la hipótesis nula.

c) $FB = \{P(H_0|x)/[1 - P(H_0|x)]\} \div \{P(H_0)/[1 - P(H_0)]\} = (0.04/0.96) \div (0.002/0.998) \approx 20.79$. No se puede rechazar la hipótesis nula.

11.83. En el problema 11.21, determinar qué se puede concluir si se utiliza el criterio del factor de Bayes: rechazar H_0 si $FB < 1$.

Puesto que la distribución a priori en este problema es impropia, la razón de posibilidades a priori no está definida. Por tanto, el criterio del factor de Bayes no se puede usar en este caso.

11.84. Si en el ejemplo 11.18 una segunda muestra de 100 observaciones produce un tiempo de reacción medio de 0.25 seg. Probar la hipótesis nula $H_0 : \theta \geq 0.3$ contra la alternativa $H_1 : \theta < 0.3$ utilizando el criterio del factor de Bayes para rechazar la hipótesis nula si $FB < 0.05$.

Se considera que la distribución a priori de θ es la distribución a posteriori que se obtuvo en el ejemplo 11.18: normal con media de 0.352 y varianza de 0.004. Al aplicar el teorema 11-3 con esta a priori y los nuevos datos, se obtienen una media de 0.269 y una varianza de 0.0007, ambas a posteriori. Con base en ello, se obtiene la probabilidad a posteriori de la hipótesis nula como 0.12.

Observamos que la probabilidad a priori de la hipótesis nula que se necesita para calcular el factor de Bayes en este problema debe basarse en la distribución a priori que se dio en el ejemplo 11.18: normal con media de 0.4 y varianza de 0.13. Si utilizamos esta última, obtenemos la probabilidad a priori de la hipótesis nula como 0.61. El factor de Bayes es 0.087. Puesto que es mayor que 0.05, no se puede rechazar la hipótesis nula.

11.85. En el problema 11.48, probar la hipótesis nula $H_0 : \theta \geq 0.7$ contra $H_1 : \theta < 0.7$ empleando la regla del factor de Bayes para rechazar la hipótesis nula si $FB < 1$.

La distribución a priori de θ es gamma con los parámetros 0.16 y 2.5. La distribución a posteriori de θ es gamma con los parámetros 10.16 y 0.04. Las probabilidades a priori y a posteriori de la hipótesis nula son, respectivamente, 0.154 y 0.022. Puesto que $P(H_0|\boldsymbol{x}) \leq P(H_0)$ se rechaza la hipótesis nula (vea el problema 11.79).

DISTRIBUCIONES PREDICTIVAS BAYESIANAS

11.86. La variable aleatoria X tiene una distribución de Bernoulli con probabilidad de éxito θ, la cual tiene una distribución a priori beta con los parámetros $\alpha = \beta = 2$. Una muestra de n ensayos en X produjo n éxitos. Se considera realizar una muestra y dos ensayos futuros. Determinar *a*) la distribución predictiva del número de éxitos futuros y *b*) la media predictiva.

a) $P(Y = y|\theta) = \binom{2}{y}\theta^y(1 - \theta)^{2-y}$, $y = 0, 1, 2$, y según el teorema 11-1, $\pi(\theta \mid \boldsymbol{x})$ es beta con parámetros $\alpha = n + 2$ y $\beta = 2$.

$f^*(y) = \displaystyle\int_0^1 \binom{2}{y}\theta^y(1 - \theta)^{2-y}\frac{\theta^{n+1}(1 - \theta)}{B(n + 2, 2)}d\theta = \binom{2}{y}\frac{B(y + n + 2, 4 - y)}{B(n + 2, 2)}$, $y = 0, 1, 2$ que se tabula como sigue:

Tabla 11-9

y	0	1	2
$f^*(y)$	$6/[(n+4)(n+5)]$	$4(n+2)/[(n+4)(n+5)]$	$[(n+2)(n+3)]/[(n+4)(n+5)]$

b) La media predictiva es $(2n^2 + 14n + 20)/[(n+4)\,(n+5)]$.

11.87. X es una variable aleatoria de Poisson con parámetro λ. Con una muestra inicial de tamaño n se obtiene una λ con una distribución a posteriori gamma y parámetros $n\bar{x} + \alpha$ y $\beta/(1 + n\beta)$. Se planea hacer una observación adicional en la población original. *a*) Calcular la distribución predictiva de la observación adicional. *b*) Demostrar que la media predictiva es igual a la media a posteriori.

a) Sea Y la observación futura. La densidad predictiva de Y está dada por

$$f^*(y) \propto \int_0^\infty \frac{e^{-\lambda}\lambda^y}{y!} \cdot \frac{(1+n\beta)^{n\bar{x}+\alpha}\lambda^{(n\bar{x}+\alpha)-1}e^{-\lambda(n\beta+1)/\beta}}{\beta^{n\bar{x}+\alpha}\Gamma(nx+\alpha)}\,d\lambda$$

$$\propto \int_0^\infty \frac{(1+n\beta)^{n\bar{x}+\alpha}\lambda^{(n\bar{x}+y+\alpha)-1}e^{-\lambda(n\beta+\beta+1)/\beta}}{y!\beta^{n\bar{x}+\alpha}\Gamma(n\bar{x}+\alpha)}\,d\lambda$$

$$= \frac{(1+n\beta)^{n\bar{x}+\alpha}}{y!\beta^{n\bar{x}+\alpha}\Gamma(n\bar{x}+\alpha)} \cdot \frac{\beta^{n\bar{x}+y+\alpha}\Gamma(n\bar{x}+y+\alpha)}{(n\beta+\beta+1)^{n\bar{x}+y+\alpha}}$$

$$f^*(y) = \binom{n\bar{x}+y+\alpha-1}{n\bar{x}+\alpha-1}\left(\frac{n\beta+1}{n\beta+\beta+1}\right)^{n\bar{x}+\alpha}\left(\frac{\beta}{n\beta+\beta+1}\right)^y, y = 0, 1, \ldots$$

Con $u = n\bar{x} + \alpha + y$ es posible escribir el lado derecho de la ecuación anterior como

$$\binom{u-1}{(n\bar{x}+\alpha)-1}\left(\frac{n\beta+1}{n\beta+\beta+1}\right)^{n\bar{x}+\alpha}\left(\frac{\beta}{n\beta+\beta+1}\right)^{u-(n\bar{x}+\alpha)}, u = n\bar{x}+\alpha, (n\bar{x}+\alpha)+1, \ldots$$

que es una función de probabilidad binomial negativa con los parámetros $r = n\bar{x} + \alpha$ y $p = \dfrac{n\beta+1}{n\beta+\beta+1}$.

b) La media de la distribución anterior es $\dfrac{r}{p} = \dfrac{(n\bar{x}+\alpha)(n\beta+\beta+1)}{(n\beta+1)}$. La media predictiva de Y resulta entonces igual a

$$\frac{(n\bar{x}+\alpha)(n\beta+\beta+1)}{(n\beta+1)} - (n\bar{x}+\alpha) = \frac{\beta(n\bar{x}+\alpha)}{(n\beta+1)}$$

que es la misma que la media a posteriori de λ.

11.88. En el problema 11.21, determinar la distribución de la media de una muestra futura de tamaño m.

$f^*(y) = \displaystyle\int_0^\infty f(y|\lambda)\pi(\lambda|\boldsymbol{x})\,d\lambda \propto \frac{1}{\Pi y_i!}\int_0^\infty e^{-(n+m)\lambda}\lambda^{n\bar{x}+m\bar{y}-\frac{1}{2}}\,d\lambda$, $\lambda > 0$. Normalizando esta densidad gamma se tiene

$$f^*(\bar{y}) = \frac{n^{n\bar{x}+\frac{1}{2}}\Gamma\left(n\bar{x}+m\bar{y}+\frac{1}{2}\right)}{\Pi y_i!\Gamma\left(n\bar{x}+\frac{1}{2}\right)(n+1)^{n\bar{x}+m\bar{y}+\frac{1}{2}}}, \bar{y} = \frac{0}{m}, \frac{1}{m}, \frac{2}{m}, \ldots.$$

11.89. Se sabe que el número de accidentes por mes en cierto tramo de una carretera sigue la distribución de Poisson con media λ. Durante los últimos 10 meses en ese tramo ocurrieron un total de 24 accidentes. ¿Cuáles son las probabilidades de que ocurran ahí más de tres accidentes durante el siguiente mes? Se considera una a priori tipo Jeffrey para λ: $\pi(\lambda) = 1/\sqrt{\lambda}$, $\lambda > 0$.

La distribución predictiva del número de accidentes Y durante el siguiente mes se puede obtener del problema 11.88 con $n = 10$, $n\bar{x} = 24$, $m = 1$:

$$f^*(y) = \frac{10^{24+\frac{1}{2}}\Gamma\left(24 + y + \frac{1}{2}\right)}{y!\Gamma\left(24 + \frac{1}{2}\right)11^{24+y+\frac{1}{2}}}, \; y = 0, 1, 2, \ldots .$$

La probabilidad que se necesita es $1 - [f^*(0) + f^*(1) + f^*(2) + f^*(3)] = 1 - [0.097 + 0.216 + 0.250 + 0.201] = 0.236$.

11.90. En el problema 11.65, ¿cuáles son las probabilidades de que el número de reacciones adversas durante el próximo año no excedan de 1?

Se necesita la distribución predictiva para una observación futura. Se tiene la distribución a posteriori del problema 11.65 como gamma con los parámetros 20.5 y 0.1. Si combinamos todo ello con la función de probabilidad de Y, obtenemos

$$f(y; \lambda) = f(y|\lambda)\pi(\lambda|\boldsymbol{x}) = \frac{e^{-\lambda}\lambda^y}{y!} \cdot \frac{10^{20.5}\lambda^{19.5}e^{-10\lambda}}{\Gamma(20.5)}, \; y = 0, 1, 2, \ldots ; \; \lambda > 0.$$

La función de probabilidad marginal para Y, que se obtuvo al integrarse respecto a λ, es

$$f^*(y) = \int_0^\infty \frac{10^{20.5}\lambda^{y+19.5}e^{-11\lambda}}{y!\Gamma(20.5)}\, d\lambda = \frac{10^{20.5}\Gamma(y + 20.5)}{y!\Gamma(20.5)11^{y+20.5}}$$

Las probabilidades correspondientes a los valores y de 0 a 7 se indican en la tabla 11-10. La probabilidad de que el número de reacciones adversas fuera 0 o bien 1 es 0.4058.

Tabla 11-10

y	0	1	2	3	4	5	6	7
$f^*(y)$	0.1417	0.2641	0.2581	0.1760	0.0940	0.0419	0.0162	0.0056

11.91. En el teorema 11-4, se supone que otra muestra independiente de tamaño 1 se toma de la población exponencial. *a*) Determinar su distribución predictiva. *b*) Estimar el resultado de la observación futura con base en la media predictiva.

a) Denotamos la observación futura por medio de Y. En consecuencia, tenemos la siguiente distribución conjunta de Y y la densidad a posteriori de θ.

$$f(y; \theta) = f(y|\theta)\pi(\theta|\,\boldsymbol{x}) = \frac{\theta e^{-\theta y}(1 + n\beta\bar{x})^{n+\alpha}\theta^{n+\alpha-1}e^{-\theta\left(\frac{1}{\beta}+n\bar{x}\right)}}{\beta^{n+\alpha}\Gamma(n + \alpha)} = \frac{(1 + n\beta\bar{x})^{n+\alpha}\theta^{n+\alpha}e^{-\theta\left(\frac{1}{\beta}+n\bar{x}+y\right)}}{\beta^{n+\alpha}\Gamma(n + \alpha)}, \quad \theta > 0.$$

Integrando respecto a θ,

$$f^*(y) = \int_0^\infty \frac{(1 + n\beta\bar{x})^{n+\alpha}\theta^{n+\alpha}e^{-\theta\left(\frac{1}{\beta}+n\bar{x}+y\right)}}{\beta^{n+\alpha}\Gamma(n + \alpha)}\, d\theta = \frac{(1 + n\beta\bar{x})^{n+\alpha}\beta^{n+\alpha+1}\Gamma(n + \alpha + 1)}{(1 + n\beta\bar{x} + \beta y)^{n+\alpha+1}\beta^{n+\alpha}\Gamma(n + \alpha)}$$

$$= \frac{(1 + n\beta\bar{x})^{n+\alpha}\beta(n + \alpha)}{(1 + n\beta\bar{x} + \beta y)^{n+\alpha+1}}, \; y > 0$$

b) La media de la distribución predictiva anterior es $\displaystyle\int_0^\infty \frac{y(1 + n\beta\bar{x})^{n+\alpha}\beta(n + \alpha)}{(1 + n\beta\bar{x} + \beta y)^{n+\alpha+1}}\, dy = \frac{1 + n\beta\bar{x}}{\beta(n + \alpha - 1)}$.

11.92. En el problema 11.29, calcular la densidad y la media predictivas de una observación futura.

$$f^*(y) = \binom{y-1}{r-1}\frac{1}{B(\alpha + nr, \beta + n\bar{x} - nr)}\int_0^1 \theta^{\alpha+nr+r-1}(1-\theta)^{\beta+n\bar{x}+y-nr-r-1}d\theta,\ y = r, r+1, \ldots$$

$$= \binom{y-1}{r-1}\frac{B(\alpha + nr + r, \beta + n\bar{x} + y - nr - r)}{B(\alpha + nr, \beta + n\bar{x} - nr)},\ y = r, r+1, \ldots$$

11.93. Una pareja tiene dos hijos y los dos son autistas. Determinar la probabilidad de que su próximo hijo también será autista bajo el supuesto de que la incidencia de autismo es independiente entre un niño y el otro, y que tiene la misma probabilidad θ. Suponga que la distribución a priori de θ es *a*) uniforme, *b*) beta con los parámetros $\alpha = 2$, $\beta = 3$.

a) Al aplicar el teorema 11-11 con $n = 2$, $x = 2$, $m = 1$ y $\alpha = \beta = 1$, se observa que la distribución predictiva de Y es $f^*(y) = \dfrac{B(3 + y, 2 - y)}{B(3, 1)} = \dfrac{(2 + y)!(1 - y)!}{8}$, $y = 0,1$. La probabilidad de que el próximo hijo sea autista es de $3/4$.

b) Con base en el teorema 11-11, con $n = 2$, $x = 2$, $m = 1$ y $\alpha = 2$, $\beta = 3$, se ve que la distribución predictiva de Y es $f^*(y) = \dfrac{B(4 + y, 4 - y)}{B(4, 3)} = \dfrac{(3 + y)!(3 - y)!}{84}$, $y = 0, 1$. La probabilidad de que el próximo hijo sea autista es de $4/7$.

11.94. Una muestra aleatoria de tamaño 20 de una población normal con media θ desconocida y varianza 4 produce una media muestral de 37.5. La distribución a priori de θ es normal con media 30 y varianza 5. Después se hace otra observación independiente de la misma población. Calcular *a*) la probabilidad predictiva de que la observación anterior no exceda de 37.5 y *b*) el intervalo predictivo de áreas de colas iguales a 95% para la observación. Según el teorema 11-12, la densidad predictiva es normal con media 37.21 y desviación estándar 2.05.

a) Igual al área de la izquierda de 0.14 bajo la densidad predictiva: 0.56.

b) $37.21 \pm (1.96 \times 2.05) = [33.19, 41.23]$.

11.95. Se lanza 10 veces una moneda y todas resultan caras. Se supone que la densidad a priori de la probabilidad de caras es $\pi(\theta) = 6\theta^5$, $0 < \theta < 1$. Determinar *a*) la distribución predictiva del número de caras en cuatro lanzamientos futuros, *b*) la media predictiva y *c*) la moda predictiva.

a) Observemos que la densidad a priori es beta con los parámetros $\alpha = 6$ y $\beta = 1$. A partir de (19), con $m = 10$, $n = 4$, $\alpha = 6$, $\beta = 1$ y $x = 10$, obtenemos $f^*(y) = \binom{4}{y}\dfrac{B(16 + y, 5 - y)}{B(16, 1)}$, $y = 0, 1, 2, 3, 4$. Los valores numéricos se muestran en la tabla 11-11.

Tabla 11-11

y	0	1	2	3	4
$f^*(y)$	0.0002	0.0033	0.0281	0.1684	0.8000

b) la media predictiva es 3.76.

c) la moda predictiva es 4.

11.96. Demostrar el teorema 11-12.

Puesto que $\overline{Y}$ es normal con media θ y varianza σ^2/m, la demostración es casi la misma que la del caso $m = 1$ con σ^2 sustituida por σ^2/m. Ello se demuestra como sigue:

La densidad predictiva $f^*(\bar{y})$ de $\overline{Y}$ está dada por

$$f^*(\bar{y}) = f(\bar{y}|\boldsymbol{x}) = \int f(\bar{y}, \theta|\boldsymbol{x})\,d\theta = \int f(\bar{y}|\theta)\cdot\pi(\theta|\boldsymbol{x})\,d\theta \propto \int_{-\infty}^{\infty} e^{-\frac{m}{2\sigma^2}(\bar{y}-\theta)^2}e^{-\frac{1}{2v_{\text{post}}}(\theta-\mu_{\text{post}})^2}d\theta$$

Después de simplificar, obtenemos

$$f^*(\bar{y}) = \int_{-\infty}^{\infty} e^{-\frac{1}{2(\sigma^2 v_{\text{post}}^2)/(\sigma^2+mv_{\text{post}}^2)}\left[\theta-\frac{(mv_{\text{post}}^2\bar{y}+\alpha^2\mu_{\text{post}})}{\sigma^2+mv_{\text{post}}^2}\right]^2} \times e^{-\frac{1}{2(\sigma^2 v_{\text{post}}^2)/(\sigma^2+mv_{\text{post}}^2)}\left[-\left(\frac{mv_{\text{post}}^2\bar{y}+\sigma^2\mu_{\text{post}}}{\sigma^2+mv_{\text{post}}^2}\right)^2+\frac{mv_{\text{post}}^2\bar{y}+\sigma^2\mu_{\text{post}}}{\sigma^2+mv_{\text{post}}^2}\right]} d\theta$$

Es posible simplificar más el exponente del segundo factor y obtener

$$f^*(y) \propto \int_{-\infty}^{\infty} e^{-\frac{1}{2(\sigma^2 v_{\text{post}}^2)/(\sigma^2+uv_{\text{post}}^2)}\left[\theta-\frac{(uv_{\text{post}}^2\bar{y}+\sigma^2\mu_{\text{post}})}{\sigma^2+mv_{\text{post}}^2}\right]^2} \times e^{-\frac{m}{2(\sigma^2+mv_{\text{post}}^2)}(\bar{y}-\mu_{\text{post}})^2} d\theta$$

El segundo factor aquí es independiente de θ. El primer factor es una densidad normal en θ y se integra en una expresión que no depende de θ ni de $\bar{y}$. Por tanto, tenemos la siguiente densidad predictiva normal para $\overline{Y}$:

$$f^*(\bar{y}) \propto e^{-\frac{m}{2(\sigma^2+mv_{\text{post}}^2)}(\bar{y}-\mu_{\text{post}})^2}$$

11.97. La variable aleatoria X tiene una distribución binomial con $n = 6$ y probabilidad de éxito θ desconocida que tiene la a priori tipo Haldane $\pi(\theta) = \dfrac{1}{\theta(1-\theta)}$, $0 < \theta < 1$. De una observación en X se obtienen tres éxitos. Si se hace otra observación en X, ¿cuántos éxitos pueden esperarse?

La distribución predictiva del número de éxitos en la segunda observación se puede obtener del teorema 11-11 (con $m = n = 6$, $x = 3$, $\alpha = \beta = 0$) como

$$f^*(y) = \binom{6}{y}\frac{B(3+y, 9-y)}{B(3,3)} \quad y = 0, 1, \ldots, 6$$

Este resultado se muestra en la tabla 11-12.

Tabla 11-12

y	0	1	2	3	4	5	6
$f^*(y)$	0.0606	0.1364	0.1948	0.2165	0.1948	0.1364	0.0606

La esperanza de esta distribución es 3. En consecuencia, es posible esperar tres éxitos en seis ensayos futuros.

PROBLEMAS DIVERSOS

11.98. Demostrar que la estimación de máxima verosimilitud de α en la distribución exponencial (vea la página 124) es $1/\overline{x}$.

Se tiene $L = \alpha^n e^{-\alpha\sum x_k}$. Por tanto, $\ln L = n \ln \alpha - \alpha\sum x_k$. Diferenciando con respecto a α e igualando a 0 resulta $\dfrac{n}{\alpha} - \sum x_k = 0$ o bien $\alpha = \dfrac{n}{\sum x_k} = \dfrac{1}{\overline{x}}$.

11.99. La variable aleatoria X tiene una distribución gamma con parámetros α y β. Demostrar que $Y = 1/X$ tiene la *densidad gamma inversa* con parámetros α y β, definida por

$$g(y) = \begin{cases} \dfrac{\beta^\alpha y^{-\alpha-1}e^{-\beta/y}}{\Gamma(\alpha)} & y > 0 \\ 0, & y \le 0 \end{cases} \quad (\alpha, \beta > 0)$$

A partir de (33), capítulo 2, tenemos

$$g(y) = \frac{(1/y)^{\alpha-1}e^{-1/(\beta y)}}{\beta^\alpha\Gamma(\alpha)} \cdot \frac{1}{y^2} = \frac{y^{-\alpha-1}e^{-\frac{1}{\beta y}}}{\beta^\alpha\Gamma(\alpha)} \quad y > 0$$

La media, la moda y la varianza son:

$$\text{Media} = \frac{\beta}{\alpha-1} \text{ para } \alpha > 1, \quad \text{Moda} = \frac{\beta}{\alpha+1}, \quad \text{Varianza} = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} \text{ para } \alpha > 2.$$

11.100. Demostrar que la estimación de Bayes con la función de pérdida del error absoluto es la mediana a posteriori. Se supone que la distribución a posteriori es continua (vea la página 83).

Es necesario demostrar que si m es la mediana de la densidad a posteriori $\pi(\theta \mid x)$, entonces

$$\int_{-\infty}^{\infty} |\theta - m|\,\pi(\theta|x)\,d\theta \le \int_{-\infty}^{\infty} |\theta - a|\,\pi(\theta|x)\,d\theta \text{ para todo } a.$$

Suponemos que $a \le m$.

$$\int_{-\infty}^{\infty} (|x - m| - |x - a|)f(x)\,dx = \int_{-\infty}^{a} (m - a)\pi(\theta|x)\,d\theta + \int_{a}^{m} (m + a - 2x)\pi(\theta|x)\,d\theta + \int_{m}^{\infty} (a - m)\pi(\theta|x)\,d\theta$$

$$\le \int_{-\infty}^{a} (m - a)\pi(\theta|x)\,d\theta + \int_{a}^{m} (m - a)\pi(\theta|x)\,d\theta + \int_{m}^{\infty} (a - m)\pi(\theta|x)\,d\theta$$

(puesto que, en la integral que se encuentra en la parte intermedia, $m + a - 2x = (m - x) - (x - a) \le m - x \le m - a$)

$$\le (m - a)\left\{\int_{-\infty}^{m} \pi(\theta|x)\,d\theta - \int_{m}^{\infty} \pi(\theta|x)\,d\theta\right\} = 0$$

La prueba cuando $a > m$ es parecida.

11.101. Generalizar los resultados del problema 11.91 a la media muestral de m observaciones futuras.

a) Denotamos la media de la muestra futura de tamaño m por $\overline{Y}$. En consecuencia, tenemos la siguiente distribución conjunta de $\overline{Y}$ y la densidad a posteriori de θ.

$$f(\bar{y};\theta) = f(\bar{y}|\theta)\pi(\theta|\boldsymbol{x}) = \frac{\theta^m e^{-\theta m\bar{y}}(1 + n\beta\bar{x})^{n+\alpha}\theta^{n+\alpha-1}e^{-\theta\left(\frac{1}{\beta}+n\bar{x}\right)}}{\beta^{n+\alpha}\Gamma(n + \alpha)} = \frac{(1 + n\beta\bar{x})^{n+\alpha}\theta^{m+n+\alpha}e^{-\theta\left(\frac{1}{\beta}+n\bar{x}+m\bar{y}\right)}}{\beta^{n+\alpha}\Gamma(n + \alpha)} \quad \theta > 0$$

Integrando con respecto a θ, tenemos

$$f^*(\bar{y}) = \int_0^{\infty} \frac{(1 + n\beta\bar{x})^{n+\alpha}\theta^{m+n+\alpha}e^{-\theta\left(\frac{1}{\beta}+n\bar{x}+m\bar{y}\right)}}{\beta^{n+\alpha}\Gamma(n + \alpha)}d\theta = \frac{(1 + n\beta\bar{x})^{n+\alpha}\beta^{m+n+\alpha}\Gamma(m + n + \alpha)}{(1 + n\beta\bar{x} + m\beta\bar{y})^{m+n+\alpha}\beta^{n+\alpha}\Gamma(n + \alpha)}$$

$$= \frac{(1 + n\beta\bar{x})^{n+\alpha}\beta^{m}\Gamma(m + n + \alpha)}{(1 + n\beta\bar{x} + m\beta\bar{y})^{m+n+\alpha}\Gamma(n + \alpha)} \qquad \bar{y} > 0$$

b) La media de la distribución anterior es

$$\int_0^{\infty} \bar{y} \cdot \frac{(1 + n\beta\bar{x})^{n+\alpha}\beta^{m}\Gamma(m + n + \alpha)}{(1 + n\beta\bar{x} + m\beta\bar{y})^{m+n+\alpha}\Gamma(n + \alpha)}d\bar{y} = \frac{\Gamma(m + n + \alpha - 2)}{m^2\Gamma(n + \alpha)}\left(\frac{\beta}{1 + n\beta\bar{x}}\right)^{m-2}$$

PROBLEMAS SUPLEMENTARIOS

PROBABILIDAD SUBJETIVA

11.102. Identifique el tipo de la probabilidad que se utiliza en los siguientes casos: *a*) No tengo idea si aprobaré o no este examen, así que diría que estoy 50% seguro de aprobarlo. *b*) Las probabilidades son dos en cinco de que tome una moneda de diez centavos porque sé que la caja tiene dos monedas de diez y tres de cinco centavos. *c*) Con base en sus marcas, hay 80% de posibilidades de que encestará más de 40 canastas que en el juego de mañana. *d*) Existe una posibilidad 50-50 de que usted se encuentre con un economista que piense que vamos a tener una recesión este año. *e*) Mi corredor de inversiones cree que las posibilidades son cinco en tres en cuanto a que el precio de esta acción se duplique en los siguientes dos meses.

PROBABILIDADES A PRIORI Y A POSTERIORI

11.103. Una caja contiene una moneda cargada con $P(C) = 0.2$ y una no cargada. Se elige una moneda de manera aleatoria de la caja y se lanza una vez. Si el resultado es cara, ¿cuál es la probabilidad del evento B de que la moneda seleccionada esté cargada?

11.104. La variable aleatoria X tiene una distribución de Poisson con un parámetro λ no conocido. Como se muestra en la tabla 11-13, el parámetro λ tiene la función de probabilidad a priori subjetiva, lo cual indica una ignorancia a priori. Una muestra aleatoria de tamaño 2 resulta en los valores 2 y 0 de X. Calcule la distribución a posteriori de λ.

Tabla 11-13

λ	0.5	1.0	1.5
$\pi(\lambda)$	1/3	1/3	1/3

11.105. X es una variable aleatoria binomial con n conocida y probabilidad de éxito θ desconocida. Determine la densidad a posteriori de θ si se supone una densidad a priori $\pi(\theta) = 4\theta^3,\ 0 < \theta < 1$.

MUESTREO DE UNA DISTRIBUCIÓN BINOMIAL

11.106. El número de herramientas defectuosas en cada lote de 10 que se fabrican mediante cierto proceso tiene una distribución binomial con parámetro θ. Suponga una densidad a priori vaga para θ [uniforme en (0,1)] y determine su densidad a posteriori con base en la información de que se encontraron dos herramientas defectuosas en el último lote que se inspeccionó.

11.107. En 50 lanzamientos de una moneda se obtuvieron 32 caras. Determine la distribución a posteriori de la proporción de caras θ que se obtendría en un número ilimitado de lanzamientos de la moneda. Considere una a priori no informativa [uniforme en (0,1)] para la probabilidad desconocida.

11.108. Continuando con el problema anterior, suponga que se hicieron 50 lanzamientos adicionales de la moneda y que se obtuvieron 35 caras. Calcule la última densidad a posteriori.

MUESTREO A PARTIR DE UNA DISTRIBUCIÓN DE POISSON

11.109. El número de accidentes durante un periodo de seis meses en una intersección tiene una distribución de Poisson con media λ. Se cree que λ tiene una densidad a priori gamma con parámetros $\alpha = 2$ y $\beta = 5$. Si se observó un total de 14 accidentes durante los primeros seis meses del año, calcule *a*) la densidad a posteriori, *b*) la media a posteriori y *c*) la varianza a posteriori.

11.110. El número de defectos en un carrete de hilo de 2000 metros que fabricó una máquina tiene una distribución de Poisson con media λ desconocida. La distribución a priori de λ es gamma con los parámetros $\alpha = 4$ y $\beta = 2$. Se encontró un total de 42 defectos en una muestra de 10 carretes que se examinó. Determine la densidad a posteriori de λ.

MUESTREO A PARTIR DE UNA DISTRIBUCIÓN NORMAL

11.111. De una población normal con media θ desconocida y varianza 9 se toma una muestra aleatoria de 16 observaciones. La distribución a priori de θ es normal estándar. Calcule *a*) la media a posteriori, *b*) su precisión y *c*) la precisión del estimador de máxima verosimilitud.

11.112. Se sabe que el tiempo de reacción de un individuo a cierto estímulo está distribuido de manera normal con media θ desconocida y desviación estándar conocida de 0.30 seg. Una muestra de 20 observaciones produjo un tiempo de reacción de 2 seg. Suponga que la densidad a priori de θ es normal con media de 1.5 seg y varianza $v^2 = 0.10$. Calcule la densidad a posteriori de θ.

DISTRIBUCIONES A PRIORI IMPROPIAS

11.113. La variable aleatoria X tiene la distribución de Poisson con parámetro λ. La distribución a priori de λ está dada por $\pi(\lambda) = 1/\sqrt{\lambda},\ \lambda > 0$. Una muestra aleatoria de 10 observaciones en X produce una media muestral de 3.5. Calcule la densidad a posteriori de λ.

11.114. Se sabe que una población es normal con media 0 y varianza θ desconocida. La varianza tiene la densidad a priori impropia $\pi(\theta) = 1/\sqrt{\theta},\ \theta > 0$. Si una muestra aleatoria de tamaño 5 lo forman 2.5, 3.2, 1.8, 2.1, 3.1, calcule la distribución a posteriori de θ.

DISTRIBUCIONES A PRIORI CONJUGADAS

11.115. Una muestra aleatoria de tamaño 20 que se toma de una distribución geométrica con parámetro θ (vea la página 117) produce una media de 5. La densidad a priori de θ es uniforme en el intervalo [0, 1]. Determine la distribución a posteriori de θ.

11.116. El tiempo entre llegada de clientes a un banco está distribuido de manera exponencial con media $1/\theta$, donde θ tiene una distribución gamma con parámetros $\alpha = 1$ y $\beta = 2$. Se observó que llegaron diez clientes en un periodo y se encontró que había un tiempo promedio de cinco minutos entre cada llegada. Determine la distribución a posteriori de θ.

11.117. Se sabe que una población es normal con media 0 y varianza θ no conocida. La varianza tiene la densidad a priori gamma inversa con parámetros $\alpha = 1$ y $\beta = 1$ (vea el problema 11.99). Calcule la distribución a posteriori de θ con base en la siguiente muestra aleatoria de la población: 2, 1.5, 2.5, 1.

ESTIMACIÓN PUNTUAL BAYESIANA

11.118. El tiempo de espera para sentarse en un restaurante está distribuido de manera exponencial con media $1/\theta$. La distribución a priori de θ es gamma con media de 0.1 y varianza de 0.1. Una muestra aleatoria de seis clientes tuvo un tiempo de espera promedio de 9 minutos. Calcule la estimación de Bayes para θ con la función de pérdida del *a*) error cuadrático y *b*) error absoluto.

11.119. El tiempo de vida X de un componente de computadora tiene la densidad exponencial dada por (vea la página 118) $f(x \mid \theta) = \theta e^{-\theta x}$, $x > 0$ con media $1/\theta$ desconocida. Suponga que la densidad a priori de θ es gamma con los parámetros α y β. Con base en una muestra aleatoria de n observaciones en X, calcule la estimación de Bayes de *a*) θ y *b*) $1/\theta$ con respecto a la función de pérdida del error cuadrático.

11.120. En el problema 11.29, calcule la estimación de Bayes de *a*) θ y *b*) $\theta(1 - \theta)$ con la función de pérdida del error cuadrático.

11.121. En el problema 11.33, determine la estimación de Bayes de θ con la pérdida del error cuadrático.

11.122. En el problema 11.26, calcule la estimación de Bayes con la pérdida del error cuadrático.

11.123. En el problema 11.6, parte *a*), determine la estimación de Bayes con la pérdida del error cuadrático para la varianza de la población, $n\theta(1 - \theta)$.

ESTIMACIÓN POR INTERVALO BAYESIANO

11.124. En diez ensayos de Bernoulli con la probabilidad de éxito θ se obtienen cinco éxitos. θ tiene la densidad a priori dada por $\pi(\theta) = \dfrac{1}{\theta(1 - \theta)}$, $0 < \theta < 1$. Calcule el intervalo de credibilidad bayesiano del área de colas iguales bayesiano a 90% para θ.

11.125. Una muestra aleatoria de tamaño 10 que se toma de una distribución geométrica con una probabilidad de éxito θ produce una media de 5. La densidad a priori de θ es uniforme en el intervalo [0.1]. Determine el intervalo de credibilidad del área de colas iguales a 88% para θ.

11.126. En el problema 11.30, calcule el intervalo de credibilidad bayesiano del área de dos colas iguales a 85% para θ.

11.127. En el problema 11.119, suponga que la densidad a priori de θ es gamma con los parámetros $\alpha = 0.2$ y $\beta = 0.15$. Una muestra aleatoria de 10 observaciones en X produjo un tiempo de vida promedio de siete años. Determine el intervalo de credibilidad bayesiano del área de colas iguales al 85% para *a*) θ y *b*)$1/\theta$.

PRUEBAS DE HIPÓTESIS BAYESIANAS

11.128. En el problema 11.21, suponga que una muestra de tamaño 10 produjo los valores 2, 0, 1, 1, 3, 0, 2, 4, 2, 2. Pruebe $H_0 : \lambda \leq 1$ contra $H_1 : \lambda > 1$ por medio de una prueba de Bayes de 0.05.

11.129. En el problema 11.6, suponga que $n = 50$ y $x = 14$ y pruebe la hipótesis nula $H_0 : \theta \leq 0.2$ contra la alternativa $H_1 : \theta > 0.2$ por medio de una prueba de Bayes de 0.025.

11.130. Suponga que en el ejemplo 11.18, una segunda muestra de 100 observaciones produjo un tiempo de reacción medio de 0.35 seg. Pruebe la hipótesis nula $H_0 : \theta \leq 0.3$ contra la alternativa $H_1 : \theta > 0.3$ mediante la aplicación de una prueba de Bayes de 0.05.

FACTORES DE BAYES

11.131. Se desea probar la hipótesis nula $\theta \leq 0.6$ contra la alternativa $\theta > 0.6$, donde θ es la probabilidad de éxito para un ensayo de Bernoulli. Suponga que θ tiene una distribución a priori uniforme en [0, 1] y que en 30 ensayos ocurrieron 17 éxitos. ¿Cuál es la conclusión si se decide rechazar la hipótesis nula si $FB < 1$?

11.132. El tiempo (en minutos) que un cliente del banco tiene que esperar en una cola para que lo atiendan está distribuido de manera exponencial con media $1/\theta$. La distribución a priori de θ es gamma con los parámetros $\alpha = 0.2$ y $\beta = 3$. En una muestra aleatoria de 10 clientes, la espera fue de tres minutos en promedio. Pruebe la hipótesis nula $H_0 : \theta \geq 0.7$ contra $H_1 : \theta < 0.7$ por medio de la regla del factor de Bayes para rechazar la hipótesis nula si $FB < 1$.

DISTRIBUCIONES PREDICTIVAS BAYESIANAS

11.133. En el problema 11.13, calcule la distribución y la media predictivas del número de accidentes durante los últimos seis meses del año.

11.134. Suponga que se obtuvieron 4 éxitos en 10 ensayos de Bernoulli con una probabilidad de éxito θ. Se considera llevar a cabo un conjunto independiente de 5 ensayos de Bernoulli más con la misma probabilidad de éxito. Encuentre la distribución predictiva del número de éxitos futuros. Suponga una densidad a priori uniforme para θ.

11.135. Se sabe que el número de accidentes por mes en un tramo de una carretera sigue la distribución de Poisson con parámetro λ. En el mismo tramo ocurrieron un total de 24 accidentes durante los últimos 10 meses. ¿Cuáles son las probabilidades de que se produzcan ahí menos de cuatro accidentes durante el siguiente mes? Suponga una a priori tipo Jeffrey para λ: $\pi(\lambda) = 1/\sqrt{\lambda}$, $\lambda > 0$.

11.136. Suponga que de 10 ensayos de Bernoulli todos fueron éxitos. ¿Cuáles son las probabilidades de que de cinco ensayos futuros de Bernoulli cinco sean éxitos? Suponga una densidad a priori uniforme para la probabilidad de éxito.

11.137. Una muestra de tamaño 20 de una población normal con media θ no conocida y varianza 4 produce una media muestral de 37.5. La distribución a priori de θ es normal con media 30 y varianza 3. Suponga que después se hace una observación independiente de la misma población. Calcule la probabilidad predictiva de que la observación no exceda de 37.5.

RESPUESTAS A LOS PROBLEMAS SUPLEMENTARIOS

11.102. *a*) subjetiva; *b*) clásica; *c*) frecuencia; *d*) información insuficiente: un caso igualmente convincente podría surgir de que es una probabilidad clásica, de frecuencia o subjetiva; *e*) subjetiva.

11.103. 2/7 **11.104.** La distribución a posteriori se da en la tabla 11-14.

Tabla 11-14

λ	0.5	1.0	1.5
$\pi(\lambda \mid x)$	0.42	0.41	0.17

11.105. $\pi(\theta|\boldsymbol{x}) = \dfrac{1}{B(x+4, n-x+1)}\theta^{x+3}(1-\theta)^{n-x}$, $0 < \theta < 1$.

11.106. La densidad a posteriori es beta con los parámetros 3 y 9.

11.107. La densidad a posteriori de θ es beta con $\alpha = 33$ y $\beta = 19$.

11.108. La densidad a posteriori de θ es beta con $\alpha = 68$ y $\beta = 34$.

11.109. *a*) La densidad a posteriori es gamma con los parámetros $n\bar{x} + \alpha = 14 + 2 = 16$ y $\beta/(1 + n\beta) = 5/6 \approx 0.83$.

b) la media a posteriori $= 80/6 \approx 13.33$; *c*) la varianza a posteriori $= 400/36 \approx 11.11$.

11.110. La densidad a posteriori es gamma con los parámetros $n\bar{x} + \alpha = 42 + 4 = 46$ y $\beta/(1 + n\beta) = 2/21 \approx 0.10$.

11.111. *a*) La media a posteriori de θ es $\left(\dfrac{16}{16+9}\right)\bar{x} = \dfrac{16\bar{x}}{25}$; *b*) la precisión es casi 4.34;

c) la precisión es aproximadamente 1.78.

11.112. La densidad a posteriori es normal con media 1.98 y varianza 0.0043.

11.113. La densidad a posteriori de λ es gamma con los parámetros 35.5 y 0.1 (vea el problema 11.25).

11.114. La densidad a posteriori es gamma inversa con $\alpha = 2$ y $\beta = 16.875$ (vea el problema 11.26).

11.115. La densidad a posteriori es beta con los parámetros 21 y 81 (vea el ejemplo 11.12).

11.116. La densidad a posteriori es gamma con los parámetros 11 y 0.02.

11.117. La densidad a posteriori es gamma inversa con los parámetros 3 y 7.75.

11.118. *a*) 0.11; *b*) 0.11

11.119. *a*) $E(\theta|\boldsymbol{x}) = \dfrac{\beta(\alpha + n)}{1 + n\beta\bar{x}}$; *b*) $E\left(\dfrac{1}{\theta}\middle|\boldsymbol{x}\right) = \dfrac{1 + n\beta\bar{x}}{\beta(n + \alpha - 1)}$

11.120. *a*) $\dfrac{\alpha + nr}{\alpha + \beta + n\bar{x}}$; *b*) $\dfrac{(\alpha + nr)(\beta + n\bar{x} - nr)}{(\alpha + \beta + n\bar{x} + 1)(\alpha + \beta + n\bar{x})}$

11.121. $\left(\dfrac{\sum x^2}{2} + \beta\right) \Big/ \left(\dfrac{n}{2} + \alpha - 1\right)$

11.122. $\dfrac{\sum x^2}{n - 3}$

11.123. $n \times \dfrac{B(x + 3, n - x + 2)}{B(x + 2, n - x + 1)} = \dfrac{n(x + 2)(n - x + 1)}{(n + 4)(n + 3)}$

11.124. [0.25, 0.75]

11.125. [0.13, 0.30]

11.126. [0.10, 0.28]

11.127. *a*) [0.078, 0.196]; *b*) [5.10, 12.82]

11.128. La probabilidad a posteriori de la hipótesis nula es 0.02 y puesto que es menor que 0.05 se rechaza la hipótesis nula.

11.129. *a*) La probabilidad a posteriori de la hipótesis nula es 0.04 y puesto que no es menor que 0.025, no se puede rechazar la hipótesis nula.

b) La probabilidad a posteriori de la hipótesis nula es 0.026 y puesto que no es menor que 0.025, no se puede rechazar la hipótesis nula.

c) La probabilidad a posteriori de la hipótesis nula es 0.015 y puesto que es menor que 0.025 se rechaza la hipótesis nula.

11.130. La probabilidad a posteriori de la hipótesis nula es 0.03 y puesto que es menor que 0.05 se rechaza la hipótesis nula.

11.131. La razón de posibilidades a posteriori es $0.66/0.34 = 1.94$ y la a priori es $6/4 = 1.5$. $FB = 1.29$. No se puede rechazar la hipótesis nula.

11.132. Se debe rechazar la hipótesis nula puesto que la probabilidad a posteriori de la hipótesis nula es 0.033, mientras que la probabilidad a priori de la hipótesis nula es 0.216.

11.133. $f^*(y) = \dbinom{y + 9}{9}\left(\dfrac{6}{11}\right)^{10}\left(\dfrac{5}{11}\right)^{y}$, $y = 0, 1, 2, \ldots$ Media predictiva $= 50/6$.

11.134.

Tabla 11-15

y	0	1	2	3	4	5
$f^*(y)$	0.106	0.240	0.288	0.224	0.112	0.029

11.135. 0.764 **11.136.** 11/16 **11.137.** 0.59

PROBLEMAS APORTADOS

1. Suponga que la distribución previa para la proporción de artículos defectuosos que produce una máquina es

p	0.1	0.2
π(p)	0.6	0.4

a) Denote con x el número de artículos defectuosos entre una muestra aleatoria de tamaño 2.
b) Estime la proporción de artículos defectuosos que produce la máquina si la muestra aleatoria de tamaño 2 produce 2 artículos defectuosos.

2. Una empresa de equipo eléctrico fabrica bombillas de luz que tienen una duración que está distribuida de forma aproximadamente normal con una desviación estándar de 100 horas. La experiencia anterior indica que μ es un valor de una variable aleatoria normal con una media $\mu_0 = 800$ horas y una desviación estándar $\sigma_0 = 10$ horas. Si una muestra aleatoria de 25 bombillas tiene una duración promedio de 780 horas, encuentre un intervalo bayesiano de 95% para μ.

3. El departamento de física de una universidad diseña un examen para alumnos de nuevo ingreso. Los miembros del departamento de física consideran que la calificación promedio para este examen variará de un grupo a otro de los alumnos de nuevo ingreso. Esta variación de la calificación promedio del grupo se expresa de manera subjetiva mediante una distribución normal, con una media $\mu_0 = 72$ y desviación estándar $\mu_0 = 2.4$.

a) ¿Qué probabilidad a priori de que la calificación promedio real, que asigna el departamento para los alumnos de nuevo ingreso del siguiente año, caiga entre 71.8 y 73.4?
b) Si el examen se aplica a una muestra aleatoria de 100 estudiantes de nuevo ingreso que tiene como resultado una calificación promedio de 70, con una desviación estándar de 8, construya un intervalo bayesiano de 95% para μ.
c) ¿Qué probabilidad a posteriori debería asignar el departamento al evento en el inciso *a*)?

APÉNDICE A

Temas matemáticos

SUMAS ESPECIALES

Las siguientes son sumas de series que surgen en la práctica. Por definición, 0! = 1. Donde la serie es infinita se indica el rango de convergencia.

1. $\displaystyle\sum_{j=1}^{m} j = 1 + 2 + 3 + \cdots + m = \frac{m(m+1)}{2}$

2. $\displaystyle\sum_{j=1}^{m} j^2 = 1^2 + 2^2 + 3^2 + \cdots + m^2 = \frac{m(m+1)(2m+1)}{6}$

3. $\displaystyle e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{j=0}^{\infty} \frac{x^j}{j!}$ toda x

4. $\displaystyle \operatorname{sen} x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{j=0}^{\infty} \frac{(-1)^j x^{2j+1}}{(2j+1)!}$ toda x

5. $\displaystyle \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{j=0}^{\infty} \frac{(-1)^j x^{2j}}{(2j)!}$ toda x

6. $\displaystyle \frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots = \sum_{j=0}^{\infty} x^j \qquad |x| < 1$

7. $\displaystyle \ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots = -\sum_{j=1}^{\infty} \frac{x^j}{j} \qquad -1 \le x < 1$

FÓRMULAS DE EULER

8. $e^{i\theta} = \cos\theta + i \operatorname{sen}\theta, \qquad e^{-i\theta} = \cos\theta - i \operatorname{sen}\theta$

9. $\displaystyle \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \qquad \operatorname{sen}\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$

FUNCIÓN GAMMA

La *función gamma*, que se denota por $\Gamma(n)$, se define mediante

$$\Gamma(n) = \int_0^{\infty} t^{n-1} e^{-t}\, dt \qquad n > 0 \tag{1}$$

Una *fórmula de recurrencia* está dada por

$$\Gamma(n+1) = n\Gamma(n) \tag{2}$$

donde $\Gamma(1) = 1$. Una extensión de la función gamma para n < 0 se puede obtener empleando (2).

Si n es un entero positivo, entonces

$$\Gamma(n+1) = n! \tag{3}$$

Por esta razón $\Gamma(n)$ algunas veces se le llama *función factorial*. Una propiedad importante de la función gamma es que

$$\Gamma(p)\Gamma(1-p) = \frac{\pi}{\operatorname{sen} p\pi} \tag{4}$$

Para $p = \frac{1}{2}$, (4) resulta

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{5}$$

Para valores grandes de n se tiene la *fórmula asintótica de Stirling*:

$$\Gamma(n + 1) \sim \sqrt{2\pi n}\, n^n e^{-n} \tag{6}$$

donde el signo $\sim$ indica que el cociente de los dos lados se acerca a 1 cuando $n \to \infty$. En particular, si n es un entero positivo grande, una buena aproximación para $n!$ está dada por

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n} \tag{7}$$

FUNCIÓN BETA

La *función beta*, que se denota mediante $B(m, n)$, se define como

$$B(m, n) = \int_0^1 u^{m-1}(1 - u)^{n-1} du \qquad m > 0, n > 0 \tag{8}$$

Ésta se relaciona con la función gamma por medio de

$$B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m + n)} \tag{9}$$

INTEGRALES ESPECIALES

Las siguientes son algunas integrales que surgen en probabilidad y la estadística.

10. $\displaystyle\int_0^\infty e^{-ax^2} dx = \frac{1}{2}\sqrt{\frac{\pi}{a}} \qquad a > 0$

11. $\displaystyle\int_0^\infty x^m e^{-ax^2} dx = \frac{\Gamma\left(\frac{m + 1}{2}\right)}{2a^{(m+1)/2}} \qquad a > 0, m > -1$

12. $\displaystyle\int_0^\infty e^{-ax^2} \cos bx\, dx = \frac{1}{2}\sqrt{\frac{\pi}{a}} e^{-b^2/4a} \qquad a > 0$

13. $\displaystyle\int_0^\infty e^{-ax} \cos bx\, dx = \frac{a}{a^2 + b^2} \qquad a > 0$

14. $\displaystyle\int_0^\infty e^{-ax} \operatorname{sen} bx\, dx = \frac{b}{a^2 + b^2} \qquad a > 0$

15. $\displaystyle\int_0^\infty x^{p-1} e^{-ax} dx = \frac{\Gamma(p)}{a^p} \qquad a > 0, p > 0$

16. $\displaystyle\int_{-\infty}^\infty e^{-(ax^2+bx+c)} dx = \sqrt{\frac{\pi}{a}} e^{(b^2-4ac)/4a} \qquad a > 0$

17. $\displaystyle\int_0^\infty e^{-(ax^2+bx+c)} dx = \frac{1}{2}\sqrt{\frac{\pi}{a}} e^{(b^2-4ac)/4a} \operatorname{fec}\left(\frac{b}{2\sqrt{a}}\right) \qquad a > 0$

donde

$$\operatorname{fec}(u) = 1 - \operatorname{fe}(u) = 1 - \frac{2}{\sqrt{\pi}}\int_0^u e^{-x^2} dx = \frac{2}{\sqrt{\pi}}\int_u^\infty e^{-x^2} dx$$

se conoce como *función de error complementaria*.

18. $\displaystyle\int_0^\infty \frac{\cos \omega x}{x^2 + a^2} dx = \frac{\pi}{2a} e^{-a\omega} \qquad a > 0, \omega > 0$

19. $\displaystyle\int_0^{\pi/2} \operatorname{sen}^{2m-1}\theta \cos^{2n-1}\theta\, d\theta = \frac{\Gamma(m)\Gamma(n)}{2\Gamma(m + n)} \qquad m > 0, n > 0$

COMBINACIÓN CONVEXA

Cuando varios puntos x_1, x_2, x_n, y escalares α_1, α_2, α_k se representan por

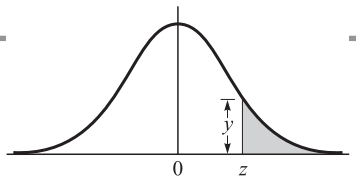
$$x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots \alpha_n x_n$$

el punto x se llama combinación lineal convexa de los puntos n dados, cuando se cumplen las condiciones

$$\lambda_1, \lambda_2, \ldots, \lambda_n, \geq 0$$

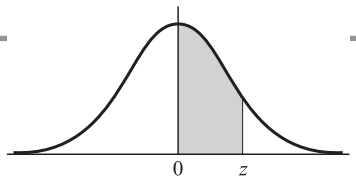
$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$$

APÉNDICE B



Ordenadas *y* de la curva normal estándar en *z*

z	0	1	2	3	4	5	6	7	8	9
0.0	.3989	.3989	.3989	.3988	.3986	.3984	.3982	.3980	.3977	.3973
0.1	.3970	.3965	.3961	.3956	.3951	.3945	.3939	.3932	.3925	.3918
0.2	.3910	.3902	.3894	.3885	.3876	.3867	.3857	.3847	.3836	.3825
0.3	.3814	.3802	.3790	.3778	.3765	.3752	.3739	.3725	.3712	.3697
0.4	.3683	.3668	.3653	.3637	.3621	.3605	.3589	.3572	.3555	.3538
0.5	.3521	.3503	.3485	.3467	.3448	.3429	.3410	.3391	.3372	.3352
0.6	.3332	.3312	.3292	.3271	.3251	.3230	.3209	.3187	.3166	.3144
0.7	.3123	.3101	.3079	.3056	.3034	.3011	.2989	.2966	.2943	.2920
0.8	.2897	.2874	.2850	.2827	.2803	.2780	.2756	.2732	.2709	.2685
0.9	.2661	.2637	.2613	.2589	.2565	.2541	.2516	.2492	.2468	.2444
1.0	.2420	.2396	.2371	.2347	.2323	.2299	.2275	.2251	.2227	.2203
1.1	.2179	.2155	.2131	.2107	.2083	.2059	.2036	.2012	.1989	.1965
1.2	.1942	.1919	.1895	.1872	.1849	.1826	.1804	.1781	.1758	.1736
1.3	.1714	.1691	.1669	.1647	.1626	.1604	.1582	.1561	.1539	.1518
1.4	.1497	.1476	.1456	.1435	.1415	.1394	.1374	.1354	.1334	.1315
1.5	.1295	.1276	.1257	.1238	.1219	.1200	.1182	.1163	.1145	.1127
1.6	.1109	.1092	.1074	.1057	.1040	.1023	.1006	.0989	.0973	.0957
1.7	.0940	.0925	.0909	.0893	.0878	.0863	.0848	.0833	.0818	.0804
1.8	.0790	.0775	.0761	.0748	.0734	.0721	.0707	.0694	.0681	.0669
1.9	.0656	.0644	.0632	.0620	.0608	.0596	.0584	.0573	.0562	.0551
2.0	.0540	.0529	.0519	.0508	.0498	.0488	.0478	.0468	.0459	.0449
2.1	.0440	.0431	.0422	.0413	.0404	.0396	.0387	.0379	.0371	.0363
2.2	.0355	.0347	.0339	.0332	.0325	.0317	.0310	.0303	.0297	.0290
2.3	.0283	.0277	.0270	.0264	.0258	.0252	.0246	.0241	.0235	.0229
2.4	.0224	.0219	.0213	.0208	.0203	.0198	.0194	.0189	.0184	.0180
2.5	.0175	.0171	.0167	.0163	.0158	.0154	.0151	.0147	.0143	.0139
2.6	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110	.0107
2.7	.0104	.0101	.0099	.0096	.0093	.0091	.0088	.0086	.0084	.0081
2.8	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063	.0061
2.9	.0060	.0058	.0056	.0055	.0053	.0051	.0050	.0048	.0047	.0046
3.0	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036	.0035	.0034
3.1	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026	.0025	.0025
3.2	.0024	.0023	.0022	.0022	.0021	.0020	.0020	.0019	.0018	.0018
3.3	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014	.0013	.0013
3.4	.0012	.0012	.0012	.0011	.0011	.0010	.0010	.0010	.0009	.0009
3.5	.0009	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0006
3.6	.0006	.0006	.0006	.0005	.0005	.0005	.0005	.0005	.0005	.0004
3.7	.0004	.0004	.0004	.0004	.0004	.0004	.0003	.0003	.0003	.0003
3.8	.0003	.0003	.0003	.0003	.0003	.0002	.0002	.0002	.0002	.0002
3.9	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001

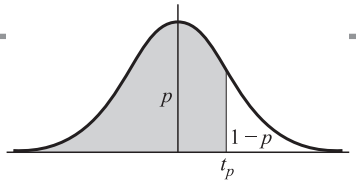


APÉNDICE C

Áreas bajo la curva normal estándar de 0 a *z*

z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

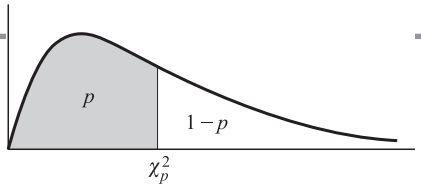
APÉNDICE D



Valores del percentil t_p para la distribución t de Student con ν grados de libertad

ν	$t_{.55}$	$t_{.60}$	$t_{.70}$	$t_{.75}$	$t_{.80}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	.158	.325	.727	1.000	1.376	3.08	6.31	12.71	31.82	63.66
2	.142	.289	.617	.816	1.061	1.89	2.92	4.30	6.96	9.92
3	.137	.277	.584	.765	.978	1.64	2.35	3.18	4.54	5.84
4	.134	.271	.569	.741	.941	1.53	2.13	2.78	3.75	4.60
5	.132	.267	.559	.727	.920	1.48	2.02	2.57	3.36	4.03
6	.131	.265	.553	.718	.906	1.44	1.94	2.45	3.14	3.71
7	.130	.263	.549	.711	.896	1.42	1.90	2.36	3.00	3.50
8	.130	.262	.546	.706	.889	1.40	1.86	2.31	2.90	3.36
9	.129	.261	.543	.703	.883	1.38	1.83	2.26	2.82	3.25
10	.129	.260	.542	.700	.879	1.37	1.81	2.23	2.76	3.17
11	.129	.260	.540	.697	.876	1.36	1.80	2.20	2.72	3.11
12	.128	.259	.539	.695	.873	1.36	1.78	2.18	2.68	3.06
13	.128	.259	.538	.694	.870	1.35	1.77	2.16	2.65	3.01
14	.128	.258	.537	.692	.868	1.34	1.76	2.14	2.62	2.98
15	.128	.258	.536	.691	.866	1.34	1.75	2.13	2.60	2.95
16	.128	.258	.535	.690	.865	1.34	1.75	2.12	2.58	2.92
17	.128	.257	.534	.689	.863	1.33	1.74	2.11	2.57	2.90
18	.127	.257	.534	.688	.862	1.33	1.73	2.10	2.55	2.88
19	.127	.257	.533	.688	.861	1.33	1.73	2.09	2.54	2.86
20	.127	.257	.533	.687	.860	1.32	1.72	2.09	2.53	2.84
21	.127	.257	.532	.686	.859	1.32	1.72	2.08	2.52	2.83
22	.127	.256	.532	.686	.858	1.32	1.72	2.07	2.51	2.82
23	.127	.256	.532	.685	.858	1.32	1.71	2.07	2.50	2.81
24	.127	.256	.531	.685	.857	1.32	1.71	2.06	2.49	2.80
25	.127	.256	.531	.684	.856	1.32	1.71	2.06	2.48	2.79
26	.127	.256	.531	.684	.856	1.32	1.71	2.06	2.48	2.78
27	.127	.256	.531	.684	.855	1.31	1.70	2.05	2.47	2.77
28	.127	.256	.530	.683	.855	1.31	1.70	2.05	2.47	2.76
29	.127	.256	.530	.683	.854	1.31	1.70	2.04	2.46	2.76
30	.127	.256	.530	.683	.854	1.31	1.70	2.04	2.46	2.75
40	.126	.255	.529	.681	.851	1.30	1.68	2.02	2.42	2.70
60	.126	.254	.527	.679	.848	1.30	1.67	2.00	2.39	2.66
120	.126	.254	.526	.677	.845	1.29	1.66	1.98	2.36	2.62
∞	.126	.253	.524	.674	.842	1.28	1.645	1.96	2.33	2.58

Fuente: R.A. Fisher y F. Yates. *Statistical tables for Biological, Agricultural and Medical Research,* publicado por Longman Group Ltd., Londres (publicado anteriormente por Oliver y Boyd, Edinburgo), y con permiso de los autores y las editoriales.

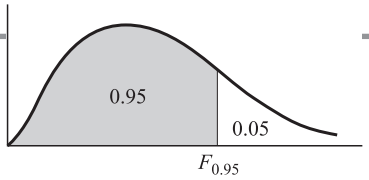


APÉNDICE E

Valores del percentil χ^2_p para la distribución ji cuadrada con *v* grados de libertad

v	$\chi^2_{.005}$	$\chi^2_{.01}$	$\chi^2_{.025}$	$\chi^2_{.05}$	$\chi^2_{.10}$	$\chi^2_{.25}$	$\chi^2_{.50}$	$\chi^2_{.75}$	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.975}$	$\chi^2_{.99}$	$\chi^2_{.995}$	$\chi^2_{.999}$
1	.0000	.0002	.0010	.0039	.0158	.102	.455	1.32	2.71	3.84	5.02	6.63	7.88	10.8
2	.0100	.0201	.0506	.103	.211	.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6	13.8
3	.0717	.115	.216	.352	.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8	16.3
4	.207	.297	.484	.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9	18.5
5	.412	.554	.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7	20.5
6	.676	.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5	22.5
7	.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3	24.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0	26.1
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6	27.9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2	29.6
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8	31.3
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3	32.9
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8	34.5
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3	36.1
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8	37.7
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3	39.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7	40.8
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2	42.3
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6	43.8
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0	45.3
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4	46.8
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8	48.3
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2	49.7
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6	51.2
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9	52.6
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	54.1
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	55.5
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	56.9
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	58.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	59.7
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8	73.4
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5	86.7
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0	99.6
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100	104	112
80	51.2	53.5	57.2	60.4	64.3	71.1	79.3	88.1	96.6	102	107	112	116	125
90	59.2	61.8	65.6	69.1	73.3	80.6	89.3	98.6	108	113	118	124	128	137
100	67.3	70.1	74.2	77.9	82.4	90.1	99.3	109	118	124	130	136	140	149

Fuente: E.S. Pearson y H.O. Hartley, *Biometric Tables for Statisticians,* vol. 1 (1966), tabla 8, pp. 137-138, con permiso.



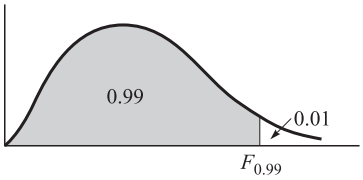
APÉNDICE F

Valores del percentil 95 (niveles de 0.05), $F_{0.95}$ para la distribución F

v_1 grados de libertad en el numerador
v_2 grados de libertad en el denominador

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Fuente: E.S. Pearson y H.O. Hartley, *Biometric Tables for Statisticians,* No. 2 (1972), tabla 5, p. 178, con permiso.



Valores del percentil 99 (niveles de 0.01), $F_{0.99}$ para la distribución F

v_1 grados de libertad en el numerador
v_2 grados de libertad en el denominador

v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4 052	5 000	5 403	5 625	5 764	5 859	5 928	5 981	6 023	6 056	6 106	6 157	6 209	6 235	6 261	6 287	6 313	6 339	6 366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.82	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Fuente: E.S. Pearson y H.O. Hartley, *Biometric Tables for Statisticians,* No. 2 (1972), tabla 5, p. 180, con permiso.

APÉNDICES G y H

Valores de $e^{-\lambda}$

$(0 < \lambda < 1)$

λ	0	1	2	3	4	5	6	7	8	9
0.0	1.0000	0.9900	0.9802	0.9704	0.9608	0.9512	0.9418	0.9324	0.9231	0.9139
0.1	0.9048	0.8958	0.8869	0.8781	0.8694	0.8607	0.8521	0.8437	0.8353	0.8270
0.2	0.8187	0.8106	0.8025	0.7945	0.7866	0.7788	0.7711	0.7634	0.7558	0.7483
0.3	0.7408	0.7334	0.7261	0.7189	0.7118	0.7047	0.6977	0.6907	0.6839	0.6771
0.4	0.6703	0.6636	0.6570	0.6505	0.6440	0.6376	0.6313	0.6250	0.6188	0.6126
0.5	0.6065	0.6005	0.5945	0.5886	0.5827	0.5770	0.5712	0.5655	0.5599	0.5543
0.6	0.5488	0.5434	0.5379	0.5326	0.5273	0.5220	0.5169	0.5117	0.5066	0.5016
0.7	0.4966	0.4916	0.4868	0.4819	0.4771	0.4724	0.4677	0.4630	0.4584	0.4538
0.8	0.4493	0.4449	0.4404	0.4360	0.4317	0.4274	0.4232	0.4190	0.4148	0.4107
0.9	0.4066	0.4025	0.3985	0.3946	0.3906	0.3867	0.3829	0.3791	0.3753	0.3716

$\lambda (= 1, 2, 3, \ldots \quad , 10)$

λ	1	2	3	4	5	6	7	8	9	10
$e^{-\lambda}$	0.36788	0.13534	0.04979	0.01832	0.006738	0.002479	0.000912	0.000335	0.000123	0.000045

Nota: Para obtener otros valores de $e^{-\lambda}$ utilizar las leyes de los exponentes.

Ejemplo: $e^{-3.48} = (e^{-3.00})(e^{-0.48}) = (.04979)(.6188) = .03081.$

Números aleatorios

51772	74640	42331	29044	46621	62898	93582	04186	19640	87056
24033	23491	83587	06568	21960	21387	76105	10863	97453	90581
45939	60173	52078	25424	11645	55870	56974	37428	93507	94271
30586	02133	75797	45406	31041	86707	12973	17169	88116	42187
03585	79353	81938	82322	96799	85659	36081	50884	14070	74950
64937	03355	95863	20790	65304	55189	00745	65253	11822	15804
15630	64759	51135	98527	62586	41889	25439	88036	24034	67283
09448	56301	57683	30277	94623	85418	68829	06652	41982	49159
21631	91157	77331	60710	52290	16835	48653	71590	16159	14676
91097	17480	29414	06829	87843	28195	27279	47152	35683	47280

Índice analítico

Ajuste de curvas, 265
Ajuste de datos mediante distribuciones teóricas, 239-241
Análisis combinatorio, 8, 17, 22, 28, 29
Análisis de varianza, 314-347
  no paramétrico, 350
  para experimentos de dos factores, 320, 330, 331
  para experimentos de tres factores, 329, 339
  para experimentos de un factor, 314, 324
  para números desiguales de observaciones, 318, 328
  para un modelo matemático lineal, 315
  tablas para, 317, 318
Análisis no paramétrico de varianza, 350
Aproximación normal de la binomial, 126, 129
Asignación de probabilidades, 6
Axiomas de probabilidad, 5

Bayesiana/o:
  distribuciones predictivas, 386, 401
  estimación puntual, 380, 394
  intervalo de estimación, 382, 397
  prueba, 383-384
  pruebas de hipótesis, 383-384, 399
Bernoulli:
  distribución de, 108
  ensayos de, 108
Bloques aleatorizados, 318, 323
Bloques en experimentos de dos factores, 318, 322
Bondad de ajuste, 219, 242, 246

Cambio de variables, 41, 42, 51, 63
Categorías, 160
Celdas, 221, 305
Centro de gravedad de los datos, 267
Centroide de los datos, 267
Clase:
  frecuencia de, 160, 176
  fronteras de, 160, 175
  intervalo de, 160, 176, 304, 305
  marca de, 160, 176
Clases y frecuencias de clase, 160, 305
Clasificación bidireccional, 318, 330
Clasificación tridireccional, 329
Clasificación unidireccional, 314
Cocientes de posibilidades a priori y a posteriori, 384
Coeficiente de:
  contingencia, 222, 250, 261
  correlación de rangos, 271, 352
  correlación linear, 289
  curtosis, 85
  determinación, 270, 301
  sesgo, 84
Coeficiente de correlación de rangos de Spearman, 271
Coeficiente de correlación generalizado, 271, 292
Coeficiente de correlación muestral, 268
Coeficientes binomiales, 10, 21
Coeficientes de correlación múltiple, 271, 293
Combinaciones, 9, 20, 29
Condicional(es):
  distribuciones, 43, 58
  esperanza, 82, 93 102
  función de la densidad, 43, 58
  función de probabilidades, 43
  momentos, 82, 93, 102
  probabilidad, 7, 14, 28
  varianza, 82, 93, 102
Confiabilidad, 195
Conjunta(s):
  distribuciones, 39, 47-51
  función de densidad, 40, 48-51
  función de distribución, 40, 47, 51
  función de probabilidad, 39, 47, 48
  tabla de probabilidades, 39
Conteo, 8, 17, 28
  principio fundamental de, 8
Convoluciones, 43, 56, 57
Corrección de Yates para continuidad, 221, 242, 244, 247
Correlación, 265-313
  coeficiente de, 82, 91, 102
  coeficiente de lineal, 270, 289
  coeficiente de múltiple, 293
  de rangos, 271, 293, 365
  fórmula de momento-producto para la, 270, 290
  generalizada, 271, 292
  interpretación probabilística de la, 295
  lineal perfecta, 268, 270
  lineal perfecta y regresión, 268, 270
  muestral, 270
  para datos agrupados, 305
  poblacional, 273
  prueba de la hipótesis para la, 274
  tabla de, 305
  teoría muestral de la, 274, 298
  y dependencia, 274
Corrida, 351
Covarianza, 81, 91, 184
Cuadrados grecolatinos, 324, 335
Cuadrados latinos, 323, 334
Curtosis, 84, 85, 96-98
Curva cuadrática, 265
Curva de aproximación, 265
Curva de mejor ajuste, 266
Curva de regresión no lineal, 271
Curva normal estándar, 110
Curva parabólica, 265
Curvas características de operación, 219, 234, 251
Curvas *CO*, 219

Deciles, 84
Decisiones estadísticas, 213
Densidad a priori de Haldane, 379
Densidades impropias, 373
Desigualdad de Chebyshev, 83, 93, 102
Desviación (en la regresión), 266
Desviación estándar, 77, 87
Diagrama de árbol, 8, 17
Diagrama de dispersión, 265, 280
Diferencias en pruebas poblacionales, 351
Diseño de experimentos, 323
Diseño experimental, 323
Distribución a priori cerrada, 379
Distribución a priori conjugada, 379, 393
Distribución a priori difusa, 373
Distribución a priori no informativa, 373
Distribución a priori vaga, 373
Distribución beta, 114, 133
Distribución bimodal, 83

Distribución binomial, 108, 118
  aproximación de Poisson de la, 128
  aproximación normal de la, 126, 129
  negativa, 117
  propiedades de, 109
  relación con la distribución de Poisson, 111
  relación con la distribución normal, 111
Distribución de Cauchy, 114, 132, 133
Distribución de *F*, 116, 138, 233
Distribución de frecuencias relativas (porcentuales), 161
Distribución de Maxwell, 118
Distribución de Pascal, 117
Distribución de Poisson, 111, 128, 129, 146
Distribución de Weibull, 118, 141
Distribución discreta, 34, 44, 45
Distribución exponencial, 118
Distribución gamma, 114, 133, 147
Distribución gaussiana, 109
Distribución geométrica, 117
Distribución hipergeométrica, 113, 131
Distribución ji-cuadrada, 115
  función generadora de momentos de la, 134
  relación con la distribución normal, 134
  relación con las distribuciones *F* y *t*, 139
  teoremas relacionados con la, 115
Distribución muestral de medias, 155, 163, 181
  cuando la varianza poblacional no se conoce, 159, 174
  teoremas relacionados cuando se conoce la varianza poblacional, 155-156
Distribución muestral, 155, 163, 166, 169, 171
  de cocientes de varianzas, 159, 174
  de diferencias y de sumas, 157, 169
  de proporciones, 156, 166
  de varianzas, 158, 169, 171
Distribución multinomial, 112, 131, 146
Distribución normal, 109, 122
  propiedades de, 110
Distribución normal bivariada, 117, 140
Distribución *t*, 115
  pruebas en las que interviene la, 236
Distribución *t* de Student, 115, 136
  pruebas en las que interviene la, 236
Distribución uniforme, 113-114, 132, 133
Distribuciones a priori impropias, 378, 392
Distribuciones a priori y a posteriori, 372, 388
  cuando se muestrea de una población binomial, 375, 390
  cuando se muestrea de una población de Poisson, 376, 391
  cuando se muestrea de una población normal, 377, 391
Distribuciones de:
  cocientes de varianzas, 174
  medias, 163
  proporciones, 166
  variaciones, 317
  varianzas, 171
Distribuciones de frecuencias porcentuales (relativas), 161
Distribuciones de probabilidad empírica, 161

Ecuaciones no lineales reducibles a la forma linear, 282
Ecuaciones normales, 267, 269
Efecto(s):
  de bloque, 320
  de la interacción, 321
  del tratamiento, 320
  residual, 321
Enfoque clásico de probabilidad, 5
Error:
  en la regresión, 266
  variación debida al, 319, 321, 322
Error estándar, 155
  tabla para el, 160
Error estándar de las estimaciones, 269, 287
Error tipo I, 213
Error tipo II, 213
Espacio muestral, 3, 10
  discreto, 4
  finito, 4
  infinito contable, 4
Esperanza matemática, 75
  algunos teoremas sobre, 76, 77
Estadística de Bose-Einstein, 31
Estadística de Fermi-Dirac, 31
Estadística no paramétrica, 184, 349-351
Estadístico, 154, 155
Estadístico de la muestra, 154
Estimación de intervalos, 195
Estimación eficiente, 195, 199
Estimación insesgada, 158, 195, 199
Estimación puntual, 195
Estimación puntual predictiva, 388
Estimador insesgado, 158, 195, 199
Estimador más eficiente, 195
Estimador sesgado, 158
Evento, 4, 10
  cierto o seguro, 4
  elemental, 4
  imposible, 4
Eventos mutuamente exclusivos, 4, 13
Expansión binomial, 108
Experimento aleatorio, 3, 10
Experimentos aleatorios, 3, 10
Experimentos de dos factores, 318, 330
Experimentos de dos factores con réplicas, 321, 331
Experimentos de Mendel, 243
Experimentos de tres factores, 329
Experimentos de un factor, 314, 324
  tablas para, 317, 318

Factor de Bayes, 384, 400
Fisher, R. A., 116, 198, 314
Fórmula de Spearman para la correlación de rangos, 352
Fórmula de Stirling, 10
Fórmula producto-momento, 270, 290

Frecuencia(s):
  de clases, 160
  distribución de, 160, 175
  enfoque de probabilidad de, 5
  histogramas de, 176
  polígonos de, 160
Frecuencia esperada, 219
Frecuencia observada, 219
Frecuencia teórica, 221
Frecuencias de las celdas, 221
Función beta, 114, 412
Función característica, 80, 90, 97
  de la distribución binomial, 109
  de la distribución de Cauchy, 132
  de la distribución de Poisson, 110
  de la distribución normal, 110
  de operación, 235
Función de densidad, 37, 38
  condicional, 43
  de la función característica, 81
  de la variable aleatoria normal estándar, 110
  de una suma de variables aleatorias, 43
  marginal, 41
  normal estándar, 110
Función de distribución:
  conjunta, 40
  marginal, 41, 48, 49
  para una variable aleatoria continua, 36, 46
  para una variable aleatoria discreta, 35, 45
  propiedades de, 35
  representación gráfica, 36, 38
Función de distribución acumulativa:
  propiedades de la, 35
  representación gráfica de la, 36, 38
Función de error [erf (z)], 110
Función de verosimilitud, 373-374
Función gamma, 114, 411
  fórmula asintótica de Stirling para, 412
Función generadora de momentos, 79, 80, 88, 96, 100
  de la distribución binomial, 121
  de la distribución de Cauchy, 132
  de la distribución de Poisson, 129
  de una suma de variables aleatorias independientes, 89
Funciones de variables aleatorias, 76

Grados de libertad:
  para la distribución *F*, 116, 321
  para la distribución ji-cuadrada, 115, 321
  para la distribución *t*, 115
Gráficas de control, 219, 238
Gráficas de control de calidad, 219, 238
Gran media, 314, 319

Hipótesis alternativa, 213
Hipótesis estadísticas, 213
Hipótesis nula, 213

Independiente(s):
  eventos, 7, 8, 14
  variables aleatorias, 41, 47, 59
  muestras, 157, 169, 197
  ensayos, 108
  variable, 265
Inferencia estadística, 153
Intervalo de confianza, 195
Intervalos de la confianza para:
  cocientes de varianzas, 198, 205
  diferencias y sumas, 197, 202, 203
  medias, 196, 200, 202
  proporciones, 197, 202, 207
  varianza, 197, 204
Intervalos de la credibilidad, 382
Intervalos predictivos, 388
Invarianza bajo transformación, 268, 271

Jacobiano, 42, 52, 54

Ley de Laplace de una sucesión, 380, 386
Ley de los grandes números, 83, 94, 103
  para ensayos de Bernoulli, 109, 122
Ley débil de los grandes números, 83
Ley fuerte de los grandes números, 83
Límites de confianza, 195
Lineal:
  coeficiente de correlación, 270, 289
  modelo para el análisis de varianza, 315
  regresión, 265, 266, 275
  relación, 265, 292

Marginal(es):
  frecuencia, 221
  funciones de densidad, 41, 58
  funciones de distribuciones, 41, 48, 49, 61
  funciones de probabilidad, 39
Máxima verosimilitud:
  estimación de la, 198, 206
  estimador de verosimilitud, 199, 206
  método de verosimilitud, 198
Media(s):
  aritmética, 75
  de funciones de variables aleatorias, 76
  de grupo, 314
  de renglón, 314
  de tratamiento, 314
  desviación, 84
  global, 314, 319
  gran, 314
  muestral, 155, 163, 177, 178
  para datos agrupados, 161, 162
Mediana, 83, 94, 95, 103
Medida de tendencia central, 75, 94, 103
Medidas de dispersión respecto a la media, 77, 96, 103
Método de mínimos cuadrados, 266
Métodos bayesianos, 372-410
Métodos de compilación, 162, 178-180, 305
Mínimos cuadrados:
  curva, 266
  curva de regresión de, 266, 272
  método de, 266
  parábola, 266, 268, 269, 284
  recta, 266-268, 275
Moda, 83, 94, 103
Momentos (*ver también r*-ésimos momentos), 78, 88, 93, 100, 177
Momentos centrales, 78
Momentos muestrales, 177
Momentos para datos agrupados, 161, 162
Muestra, 153
Muestra aleatoria, 154
Muestras emparejadas, prueba de los signos para, 348
Muestras grandes, 196, 216
Muestreo, 153
Muestreo con sustitución, 112, 153
Muestreo sin sustitución, 113, 153

Nivel de significación, 214
Niveles de confianza, 195, 214
Números aleatorios, 154
Números desiguales de observaciones, 318

Papel para representar gráficamente la curva normal, 219
Patrón cíclico, 351, 361
Patrón de tendencia, 351, 362
Percentiles, 84, 96, 103
Permutaciones, 9, 18, 28
Población(al), 153
  coeficientes de correlación, 273
  parámetros, 154
  tamaño de la, 153
Población binomial, 154
Población finita, 153, 156
Población infinita, 153, 156
Población multinomial, 220
Población normal, 154
Posibilidades, 5
Potencia de una prueba, 219, 235
Precisión de una distribución, 377
Principio fundamental de conteo, 8
Probabilidad, 5
  axiomas de, 5
  cálculo, 12, 22, 23
  concepto de, 5
  conjunta, 39, 47, 48
  de éxito, 108
  de fracaso, 108
  de un evento, 5
  discreta, 34
  distribución de, 34, 37, 44
  distribuciones de funciones de variables aleatorias, 42
  empleando análisis combinatorios, 22
  enfoque clásico de, 5
  enfoque de frecuencia de, 5
  función de densidad de, 37
  función de, 34
  interpretación de la correlación, 273, 295
  interpretación de la regresión, 271, 295
  papel para representación gráfica de la, 219
  superficie, 40
Probabilidad a posteriori, 373
Probabilidad a priori, 373
Probabilidad de intervalos, 36
Probabilidad empírica, 5
Probabilidad geométrica, 44, 60
Probabilidad subjetiva, 372, 388
Problema de la aguja de Buffon, 64
Problema del cumpleaños, 26
Problema del sobre, 25
Prueba arriba-y-debajo de la mediana, 351
Prueba de hipótesis (*ver* Pruebas de hipótesis y significancia)
Prueba de Kruskal-Wallis *H*, 283, 350, 360
  corregida por coincidencias (empates), 350
Prueba de las corridas (rachas) para el carácter aleatorio, 350, 361, 364
Prueba del signo, 348, 352
Prueba *F* para la hipótesis nula de medias iguales, 317
Prueba *H* corregida por coincidencias (empates), 350
Prueba ji-cuadrada, 219-222, 233, 242, 246, 252, 258
Prueba *U* de Mann-Whitney, 349, 354, 358
Pruebas bilaterales, 214
Pruebas de dos colas, 214
Pruebas de hipótesis y significación, 213-264.
  en las que interviene la distribución *F*, 217, 233, 257
  en las que interviene la distribución ji-cuadrada, 217, 233, 242, 257
  en las que interviene la distribución normal, 214, 222, 255
  en las que interviene la distribución *t* de Student, 218, 230
  para cocientes de varianzas, 218
  para diferencias de proporciones, 217, 227, 253
  para diferencias entre los coeficientes de correlación, 274
  para diferencias entre medias, 217, 218, 227, 255
  para el coeficiente de correlación, 274
  para el coeficiente de regresión en la regresión linear, 273
  para los valores predichos en la regresión linear, 273
  para medias, 216, 217, 222, 255
  para muestras grandes, 216, 217
  para muestras pequeñas, 217
  para proporciones, 216, 222, 255
  para varianzas, 218
Pruebas de una cola, 215
Pruebas no paramétricas, 348-371
Pruebas unilaterales, 215
Punto muestral, 3
Puntuación estándar, 78, 110

r-ésimo momento:
central, 78-79
condicional, 82
de renglón, 79
respecto de la media, 78-79
respecto del origen, 82
Rango intercuartil, 84
Rango semiintercuartil, 84
Región crítica, 214
Región de:
aceptación, 214
no significancia, 214
rechazo, 214
significancia, 214
Reglas de decisión, 213
Regresión, 265
ecuación de, 265, 269
múltiple, 269, 285
plano de, 269
recta de, 272
superficie de, 269
Relación:
entre la teoría de estimación y la prueba de hipótesis, 219
entre las distribuciones binomial y de Poisson, 111
entre las distribuciones binomial y normal, 111
entre las distribuciones de Poisson y normal, 112
entre las distribuciones ji-cuadrada, *F* y *t*, 117, 139
no lineal, 265, 292
Repeticiones, 314
Réplicas, 314, 321
Residual (en la regresión), 266
Resultados igualmente probables, 6

Serie(s) de Fourier, 81, 97
Sesgo, 84, 96, 97, 181
Significación estadística, 214
Significado de una diferencia entre coeficientes de correlación, 274
Subconjuntos medibles, 5

Tablas de contingencia, 221, 246
Tamaño de la muestra, 153
Teorema de Bayes (regla), 8, 17
Teorema del límite central 112, 129, 156, 253
para variables aleatorias binomiales, 129
prueba del, 130
Teoría de corridas (rachas), 350
Teoría de muestreo de la correlación, 274, 298
Teoría de muestreo de la regresión, 297
Transformación de Fisher en:
correlación, 274
el análisis de varianza, 314
Transformación *Z* de Fisher, 274
Transformada de Fourier, 81
Transformada inversa de Fourier, 81
Traslación de ejes, 267
Tratamientos, 314
Tratamientos y bloques, 318
Unidades estándar, 78

Valor esperado, 75, 85
algunos teoremas sobre, 76-77, 85
Valor *p*, 215
Valores críticos, 195
Valores de tendencia, 301
Valores esperados de las variaciones, 316

Variable aleatoria, 34
continua, 36, 46
discreta, 34, 44
no-discreta, 34
Variable aleatoria absolutamente continua, 36
Variable aleatoria asintóticamente normal, 111, 112, 156
Variable aleatoria distribuida normalmente, 109, 122
Variable aleatoria estandarizada, 78
Variables aleatorias dependientes, 41
Variables transformadas, 265
Variación:
dentro de los tratamientos, 315
distribución de la, 317
entre tratamientos, 315
explicada, 270, 273, 289, 292
métodos abreviados para obtener la, 315
no explicada, 270, 273, 292
para experimentos de dos factores, 319
residual, 319
total, 270, 273, 289, 315, 319
valor esperado, 316
Varianza, 77, 78, 81, 87, 100
combinada, 218
condicional, 82, 93, 102
de la distribución *F*, 116
de la distribución normal, 110
de la distribución *t* de Student, 116
de una distribución binomial, 109
distribución muestral de, 171
muestral, 157, 177
para datos agrupados, 161, 162
para muestras, 177
Verosimilitud, 198

ÍNDICE DE PROBLEMAS RESUELTOS

Ajuste de datos mediante distribuciones teóricas, 239
Aleatorias, variables (*ver* Variables aleatorias)
Análisis combinatorio, 17
  probabilidad usando, 22
Aproximación normal a la distribución binomial, 126, 129

Bayesiana/o/s:
  distribuciones predictivas, 401
  estimación por intervalo, 397
  estimación puntual, 394
  prueba de hipótesis, 399
Bondad del ajuste, 242, 246

Cálculo de probabilidades, 12
Cambio de variables, 51, 63
Clasificación bidireccional, 330
Clasificación tridireccional, 329
Clasificación unidireccional, 324
Coeficiente de:
  contingencia, 250
  correlación, 91
  correlación lineal, 289 (*ver también* Correlación)
  determinación, 301
Coeficiente de correlación generalizado, 292
Coeficiente de correlación múltiple, 293
Coeficientes binomiales, 21
Coeficientes de Fourier, 97
Combinaciones, 20
Condicional(es):
  densidad, 58
  distribución, 58
  esperanza, 93
  momentos, 93
  probabilidad, 14
  varianza, 93
Conteo, 17
Contingencia, coeficiente de, 250
Convoluciones, 56
Correlación:
  coeficiente de, 91, 140
  de rangos, 293, 365
  fórmula del producto-momento para la, 290
  generalizada, 292
  interpretación probabilística de la, 295
  lineal, 289
  múltiple, 293
  tabla de, 305
  teoría muestral de la, 298
Covarianza, 91, 184
Cuadrados grecolatinos, 335
Cuadrados latinos, 334
Curtosis, 96, 97, 98
Curva característica de operación, 234, 251

Desigualdad de Chebyshev, 93
Desviación estándar, 87
Determinación, coeficiente de, 301
Diagrama de dispersión, 280
Diagramas de árbol, 17
Dispersión, medidas de, 96
Distribución:
  a priori conjugada, 393
  a priori impropia, 392
  a priori y a posteriori, 388
  beta, 133
  condicional, 58
  conjunta, 47
  continua, 46
  de cocientes de varianzas, 174
  de diferencias y sumas, 169
  de frecuencias, 175
  de medias, 163
  de proporciones, 166
  de varianzas, 171
  de Weibull, 141
  discreta, 4
  F de Fisher, 138
  gamma, 133
  hipergeométrica, 131
  marginal, 48
  multinomial, 131
  normal, 122
  normal bivariada, 140
  predictiva Bayesiana, 401
  relación entre, *F*, ji-cuadrada y *t*, 139
  uniforme, 132
Distribución binomial, 118
  aproximación de Poisson a la, 128
  aproximación normal de la, 126, 129
  función generadora de momentos, 121
Distribución de Cauchy, 132, 133
  función característica de la, 132
  función generadora de momentos, 132
  relación con la distribución uniforme, 133
Distribución de Poisson, 128
  función generadora de momentos, 129
Distribución *F*, 138
  pruebas en las que interviene, 233
  relación con las distribuciones ji-cuadrada y *t*, 139
Distribución ji-cuadrada, 134
  función generadora de momentos de la, 134
  pruebas en las que interviene, 233, 242
  relación con la distribución normal, 134
  relación con las distribuciones *F* y *t*, 139
Distribución muestral, 163, 166, 169, 171, 174
  de la diferencia de medias, 169
  de la media cuando no se conoce la varianza poblacional, 174
  de las medias, 163, 181
  de las proporciones, 166
  de las sumas de varianzas, 169
  de las varianzas, 171
  de los cocientes de varianzas, 174
Distribución *t* de Student, 136
  pruebas en las que interviene la, 236
  relación con las distribuciones *F* y ji-cuadrada, 139
Distribuciones a priori conjugadas, 393
Distribuciones a priori y a posteriori, 388
  cuando se muestrea de una población binomial, 390
  cuando se muestrea de una población de Poisson, 391
  cuando se muestrea de una población normal, 391
Distribuciones conjuntas, 47
Distribuciones de probabilidad:
  continuas, 46
  discretas, 44

Distribuciones impropias, 392
Distribuciones teóricas, ajuste de datos mediante, 239

Ecuaciones no lineales reducibles a ecuaciones lineales, 282
Error estándar de estimación, 287
Espacios muestrales, 10
Esperanza de variables aleatorias, 85
  condicionales, 93
Estimaciones:
  de máxima verosimilitud, 206
  eficientes, 199
  insesgadas, 199
  intervalo de confianza, 200, 202
Estimaciones de máxima verosimilitud, 206
  para la media de una distribución normal, 206
  para la varianza de una distribución normal, 207
Estimaciones del intervalo de confianza para:
  cocientes de varianzas, 205
  diferencias de medias, 203
  diferencias de proporciones, 202
  la desviación estándar, 208
  la media cuando la varianza de la población no se conoce, 174
  medias en muestras grandes, 200
  medias en muestras pequeñas, 202
  proporciones, 202, 207
  varianzas, 204
Eventos, 10:
  independientes, 14
  mutuamente exclusivos, 13
Experimentos aleatorios, 10
Experimentos de dos factores, 330
  con replicación , 331
Experimentos de tres factores, 329
Experimentos de un factor, 324

Factor de Bayes, 400
Fórmula producto-momento, 290
Frecuencia:
  distribuciones de, 175
  histogramas de, 176
  polígonos de, 176
Función característica, 90, 97
  de la distribución de Cauchy, 132
  de operación, 235
Función de densidad marginal, 58
Función de distribución continua, 46
Función de distribución discreta, 45
Función de distribución marginal, 48, 61
Función generadora de momentos, 88, 96
  de sumas de variables aleatorias independientes, 89
  de una distribución binomial, 121
  de una distribución de Cauchy, 132
  de una distribución de Poisson, 129
Función potencia, 235
Funciones de densidad conjunta, 48
Funciones de distribución:
  continuas, 46
  discretas, 45
  marginales, 48, 49

Gráficas de control de calidad, 238

Interpretación probabilística de la correlación y de la regresión, 295

Ley de los grandes números, 93
  para ensayos de Bernoulli, 122

Media muestral:
  cálculo de la, 177
  fórmula de compilación para la, 178
Mediana, 94
Medias:
  cálculo de, para muestras, 177
  distribución muestral de, 163
Medidas de dispersión, 96
Medidas de tendencia central, 94
Mínimos cuadrados:
  parábola de, 284
  recta de, 275
Moda, 94
Modificaciones para números desiguales de observaciones, 328
Momentos muestrales, cálculo de los, 177
Momentos, 88
  cálculo de, 93
  condicionales, 93
  para muestras, 177

Papel log-log, 283
Patrón cíclico en una prueba de corridas (rachas), 361
Patrón de tendencias en pruebas de corridas (rachas), 362
Percentiles, 96
Permutaciones, 18
Probabilidad:
  cálculo de la, 12
  cálculo de la, utilizando análisis combinatorio, 22, 23
  condicional, 14
  distribuciones de, 44
  geométrica, 60
  subjetiva, 388
  teoremas sobre, 11
Probabilidad geométrica, aplicaciones del cálculo de la, 60
Problema de la aguja de Buffon, 64
Proporciones, distribución muestral de, 166
Prueba de las corridas (rachas) para la aleatoriedad, 361
  aplicaciones de la, 364
Prueba del signo, 352
Prueba *H* de Kruskal-Wallis, 283
Prueba ji-cuadrada de bondad del ajuste, 242, 246, 252
Prueba *U* de Mann-Whitney, 354
  distribución muestral de, 358
Pruebas:
  de las diferencias de medias, 227
  de las diferencias de proporciones, 227
  de las medias usando distribuciones normales, 222
  de las proporciones usando distribuciones normales, 222
  en las que interviene la distribución *F*, 233
  en las que interviene la distribución *t* de Student, 230.
  en que interviene la distribución ji-cuadrada, 233, 242

Regresión:
  interpretación de probabilidades de, 295
  múltiplos, 285
  recta de mínimos cuadrados, 275
  teoría muestral de, 297
Relación no lineal, 292
Relaciones lineales, 292

Serie de Fourier, 97
Sesgo, 96, 97, 181

Tablas de contingencia, 246
Tendencia central, medidas de, 94
Teorema de Bayes, 17
Teorema del límite central, 129
  para variables aleatorias binomiales, 129
  prueba del, 130
Teoría del muestreo de:
  correlación, 298
  regresión, 297

Valores de tendencia, 301
Variable aleatoria discreta, 44
Variables aleatorias:
  continuas, 46
  discretas, 44
  esperanza condicional de, 93
  esperanza de, 85
  independientes, 47, 59
Variables, cambio de, 51, 63
Variación:
  explicada, 289
  inexplicada, 292
  total, 289
Varianza 87
  cálculo de la, para muestras, 177
  condicional, 93
  distribución muestral de la, 171
Varianza muestral:
  cálculo de la, 177
  fórmula de compilación para la, 179

¡Estudie a su propio ritmo y apruebe su examen con Schaum!

Si usted no tiene mucho tiempo pero quiere sobresalir en clase, este libro le ayudará a:

- **Prepararse para los exámenes**
- **Encontrar las respuestas rápidamente**
- **Estudiar con mayor efectividad**

En www.schaums.com encontrará videos para los principales temas de la serie Schaum.

Mc Graw Hill Education

Síguenos en @MHEducacion

Encuéntranos en McGraw-Hill Educación

Visite nuestra página WEB
www.mcgraw-hill-educacion.com